

X

OM protein - protein search, using sw model

Run on: August 5, 2006, 00:05:14 ; Search time 318 Seconds  
(without alignments)  
2763.411 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_7.2:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5156	99.8	950	1 ATS15_HUMAN	Q8te58 homo sapien
2	5156	99.8	950	2 Q32MI6_HUMAN	Q32mi6 homo sapien
3	4792	92.8	950	2 Q504Z2_MOUSE	Q504z2 mus musculu
4	2916.5	56.5	882	2 Q4RYP9_TETNG	Q4ryp9 tetraodon n
5	2881	55.8	562	1 ATS15_MOUSE	P59384 mus musculu
6	2802	54.3	875	2 Q4S903_TETNG	Q4s903 tetraodon n
7	2505	48.5	947	2 Q2N1I7_PIG	Q2n1i7 sus scrofa
8	2482.5	48.1	968	1 ATS1_MOUSE	P97857 mus musculu
9	2482.5	48.1	968	2 Q3U0I4_MOUSE	Q3u0i4 mus musculu
10	2480.5	48.0	967	1 ATS1_HUMAN	Q9uhi8 homo sapien
11	2480.5	48.0	967	2 Q5HYL0_HUMAN	Q5hyl0 homo sapien
12	2472.5	47.9	928	2 Q5U261_XENLA	Q5u261 xenopus lae
13	2472.5	47.9	967	2 Q8NE26_HUMAN	Q8ne26 homo sapien
14	2471	47.9	967	1 ATS1_RAT	Q9wuq1 rattus norv
15	2466	47.8	967	2 Q68EJ2_RAT	Q68ej2 rattus norv
16	2366.5	45.8	934	2 Q4AED3_ORYLA	Q4aed3 oryzias lat
17	2307.5	44.7	759	2 Q8HZM8_HORSE	Q8hzm8 equus cabal
18	2300	44.5	911	2 Q5R6D5_PONPY	Q5r6d5 pongo pygma
19	2208.5	42.8	844	2 Q3TQF7_MOUSE	Q3tqf7 mus musculu
20	2168.5	42.0	682	2 Q3TTE6_MOUSE	Q3tte6 mus musculu
21	2135.5	41.4	890	1 ATS8_HUMAN	Q9up79 homo sapien
22	2113	40.9	905	1 ATS8_MOUSE	P57110 mus musculu
23	2099.5	40.7	901	2 Q4RYQ0_TETNG	Q4ryq0 tetraodon n
24	2032	39.3	918	2 Q5FWF1_HUMAN	Q5fwf1 homo sapien
25	1990	38.5	1906	1 ATS20_MOUSE	P59511 mus musculu
26	1980.5	38.4	2080	2 Q4SBC8_TETNG	Q4sbc8 tetraodon n
27	1963	38.0	1935	1 ATS9_HUMAN	Q9p2n4 homo sapien
28	1917	37.1	837	1 ATS4_HUMAN	O75173 homo sapien
29	1910	37.0	837	2 Q5RFQ8_PONPY	Q5rfq8 pongo pygma

30	1890.5	36.6	833	1	ATS4_MOUSE	Q8bnj2	mus	musculu
31	1890.5	36.6	845	2	Q3TNX8_MOUSE	Q3tnx8	mus	musculu
32	1888.5	36.6	845	2	Q3U025_MOUSE	Q3u025	mus	musculu
33	1888	36.6	1911	1	ATS20_HUMAN	P59510	homo	sapien
34	1861.5	36.0	867	2	Q66KM3_XENTR	Q66km3	xenopus	tro
35	1860	36.0	839	2	Q7YS95_BOVIN	Q7ys95	bos	taurus
36	1856	35.9	930	1	ATS5_MOUSE	Q9r001	mus	musculu
37	1847.5	35.8	930	1	ATS5_HUMAN	Q9una0	homo	sapien
38	1847.5	35.8	930	2	Q52LV4_HUMAN	Q52lv4	homo	sapien
39	1835	35.5	928	2	Q6TY19_RAT	Q6ty19	rattus	norv
40	1816.5	35.2	627	2	Q4RY31_TETNG	Q4ry31	tetraodon	n
41	1765.5	34.2	893	2	Q6A017_MOUSE	Q6a017	mus	musculu
42	1747.5	33.8	981	2	Q4RY30_TETNG	Q4ry30	tetraodon	n
43	1745.5	33.8	630	1	ATS4_RAT	Q9esp7	rattus	norv
44	1720.5	33.3	1322	2	Q4RSH9_TETNG	Q4rsh9	tetraodon	n
45	1685	32.6	340	2	Q91Z56_MOUSE	Q91z56	mus	musculu

# ALIGNMENTS

## RESULT 1

### ATS15\_HUMAN

ID ATS15\_HUMAN STANDARD; PRT; 950 AA.

AC Q8TE58;

DT 28-FEB-2003, integrated into UniProtKB/Swiss-Prot.

DT 01-JUN-2002, sequence version 1.

DT 07-MAR-2006, entry version 29.

DE ADAMTS-15 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15).

GN Name=ADAMTS15;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA].

RX MEDLINE=21856482; PubMed=11867212; DOI=10.1016/S0378-1119(01)00861-7;

RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,

RA Lopez-Otin C.;

RT "Cloning, expression analysis, and structural characterization of

RT seven novel human ADAMTSs, a family of metalloproteinases with

RT disintegrin and thrombospondin-1 domains.";

RL Gene 283:49-62(2002).

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted protein; extracellular space;

CC extracellular matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed in fetal liver and kidney, but not

CC in any of the adult tissues examined.

CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important

CC for a tight interaction with the extracellular matrix (By

CC similarity).

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By

CC similarity).

CC -!- SIMILARITY: Contains 1 disintegrin domain.

CC -!- SIMILARITY: Contains 1 peptidase M12B domain.

CC -!- SIMILARITY: Contains 3 TSP type-1 domains.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AJ315733; CAC86014.1; -; mRNA.

DR HSSP; P07996; 1LSL.

DR MEROPS; M12.025; -.

DR Ensembl; ENSG00000166106; Homo sapiens.

DR HGNC; HGNC:16305; ADAMTS15.

DR MIM; 607509; gene.

DR InterPro; IPR006586; ADAM\_cysteine.

DR InterPro; IPR010294; ADAM\_spacer1.

DR InterPro; IPR013273; ADAM\_TS.

DR InterPro; IPR013277; ADAM\_TS8.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01861; ADAMTS8.  
 DR PRINTS; PR01857; ADAMTSFAMILY.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Extracellular matrix; Glycoprotein; Hydrolase; Metal-binding;  
 KW Metalloprotease; Protease; Repeat; Signal; Zinc; Zymogen.  
 FT SIGNAL 1 17 Potential.  
 FT PROPEP 18 212 By similarity.  
 FT /FTId=PRO\_0000029192.  
 FT CHAIN 213 950 ADAMTS-15.  
 FT /FTId=PRO\_0000029193.  
 FT DOMAIN 218 427 Peptidase M12B.  
 FT DOMAIN 428 515 Disintegrin.  
 FT DOMAIN 516 571 TSP type-1 1.  
 FT DOMAIN 839 895 TSP type-1 2.  
 FT DOMAIN 896 949 TSP type-1 3.  
 FT REGION 701 838 Spacer.  
 FT COMBIAS 572 700 Cys-rich.  
 FT ACT\_SITE 362 362 By similarity.  
 FT METAL 361 361 Zinc (catalytic) (By similarity).  
 FT METAL 365 365 Zinc (catalytic) (By similarity).  
 FT METAL 371 371 Zinc (catalytic) (By similarity).  
 FT SITE 174 174 Cysteine switch (Potential).  
 FT CARBOHYD 141 141 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 591 591 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 623 623 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 679 679 N-linked (GlcNAc. . .) (Potential).  
 FT DISULFID 339 422 By similarity.  
 FT DISULFID 377 406 By similarity.  
 FT DISULFID 528 565 By similarity.  
 FT DISULFID 532 570 By similarity.  
 FT DISULFID 543 555 By similarity.  
 SQ SEQUENCE 950 AA; 103287 MW; 5DFBE18285CCCC3B CRC64;

Query Match 99.8%; Score 5156; DB 1; Length 950;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
 |||||  
 Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
 Qy 61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
 |||||  
 Db 61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
 Qy 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
 |||||  
 Db 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
 Qy 181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240  
 |||||  
 Db 181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN 300  
 |||||  
 Db 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN 300  
 |||||

Qy 301 KVS DKHPEYWDTAILFTRQDL CGATTCDTLGMADV GTMCDPKRSCSVIEDDGLPSAFTTA 360  
 |||||  
 Db 301 KVS DKHPEYWDTAILFTRQDL CGATTCDTLGMADV GTMCDPKRSCSVIEDDGLPSAFTTA 360  
 |||||

Qy 361 HELGHVFNMPHDNVK VCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420  
 |||||  
 Db 361 HELGHVFNMPHDNVK VCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420  
 |||||

Qy 421 DCLLDQPSKPI SLPEDLP GASYTLSQQCELA FGVGSKPCPYMQYCTKLWCTGKAKGQ MVC 480  
 |||||  
 Db 421 DCLLDQPSKPI SLPEDLP GASYTLSQQCELA FGVGSKPCPYMQYCTKLWCTGKAKGQ MVC 480  
 |||||

Qy 481 QTRHFPWADGTSCGEGKLC LKGACVERHNLNKH RVDSWAKWDPYGPCSRTC GGGVQLAR 540  
 |||||  
 Db 481 QTRHFPWADGTSCGEGKLC LKGACVERHNLNKH RVDSWAKWDPYGPCSRTC GGGVQLAR 540  
 |||||

Qy 541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600  
 |||||  
 Db 541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600  
 |||||

Qy 601 VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGCKIKAGCD 660  
 |||||  
 Db 601 VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGCKIKAGCD 660  
 |||||

Qy 661 GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720  
 |||||  
 Db 661 GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720  
 |||||

Qy 721 GDDNYLALKNSQ GKYLNGHFV VSAVERDLVVGSL LRYSGTGTA VESLQASRP ILEPLT 780  
 |||||  
 Db 721 GDDNYLALKNSQ GKYLNGHFV VSAVERDLVVGSL LRYSGTGTA VESLQASRP ILEPLT 780  
 |||||

Qy 781 VEVL SVGKMT PPRVRSFYLPKEPREDKSSH PKDPRG P SVLHNSVLSLSNQVEQPDDRPP 840  
 |||||  
 Db 781 VEVL SVGKMT PPRVRSFYLPKEPREDKSSH PKDPRG P SVLHNSVLSLSNQVEQPDDRPP 840  
 |||||

Qy 841 ARWVAGSWGPCSASC GSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900  
 |||||  
 Db 841 ARWVAGSWGPCSASC GSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900  
 |||||

Qy 901 AWS PCKSKCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC 950  
 |||||  
 Db 901 AWS PCKSKCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC 950  
 |||||

## RESULT 2

Q32MI6\_HUMAN

ID Q32MI6\_HUMAN, PRELIMINARY; PRT; 950 AA.

AC Q32MI6;

DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.

DT 06-DEC-2005, sequence version 1.

DT 07-FEB-2006, entry version 3.

DE A disintegrin-like and metalloprotease (Reprolysin type) with

DE thrombospondin type 1 motif, 15, preproprotein.

GN Name=ADAMTS15;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=PCR rescued clones;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=PCR rescued clones;

RG NIH MGC Project;

RL Submitted (NOV-2005) to the EMBL/GenBank/DDBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; BC109114; AAI09115.1; -; mRNA.

DR GO; GO:0031012; C:extracellular matrix; IEA.

DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.

DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.

DR GO; GO:0006508; P:proteolysis; IEA.

KW Integrin; Metalloprotease; Protease.

SQ SEQUENCE 950 AA; 103287 MW; 5DFBE18285CCCC3B CRC64;

Query Match 99.8%; Score 5156; DB 2; Length 950;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MLLLGILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRSRRSSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRSRRSSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPISILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSSQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSSQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC	480

Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVQGVQKCIKAGCD	660
Db	601	VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVQGVQKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Qy	721	GDDNYLALKNSQGKYLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPPILEPLT	780
Db	721	GDDNYLALKNSQGKYLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPPILEPLT	780
Qy	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCGSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC�NLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC�NLHRKPQELDFCVLRPC	950

# RESULT 3

## Q504Z2\_MOUSE

ID Q504Z2\_MOUSE PRELIMINARY; PRT; 950 AA.  
AC Q504Z2;  
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.  
DT 07-JUN-2005, sequence version 1.  
DT 07-FEB-2006, entry version 5.  
DE A disintegrin-like and metalloprotease (Reprolysin type) with  
DE thrombospondin type 1 motif, 15.  
GN Name=Adams15;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RG NIH MGC Project;  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC094677; AAH94677.1; mRNA.  
 DR Ensembl; ENSMUSG00000033453; Mus musculus.  
 DR MGI; MGI:2449569; Adamts15.  
 DR GO; GO:0031012; C:extracellular matrix; IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
 DR GO; GO:0006508; P:proteolysis; IEA.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013277; ADAM\_TS8.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002332; PII\_GlnB\_UMP\_S.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00496; PII\_GLNB\_UMP; UNKNOWN\_1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Integrin; Metalloprotease; Protease.  
 SQ SEQUENCE 950 AA; 103938 MW; 891926F281E92010 CRC64;  
  
 Query Match 92.8%; Score 4792; DB 2; Length 950;  
 Best Local Similarity 92.9%; Pred. No. 0;  
 Matches 883; Conservative 22; Mismatches 45; Indels 0; Gaps 0;

Qy	1	MLLLGILTLAGRTAGGFEPEREEVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGISILALAWRPAGSSEPEWEVVVPIRRDPDINGRHHYRRGTEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QQDFYHLHTPDAQFLAPAFATEYLGVLQRLTGSSDLRRCFYSGYVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQCAHLLQRRGVPGGPGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNTSAPEAQRHSQCAHLLQRRGAPVGPSPDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRTGAGESHNRSSRAKRFVSIPTYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLGDRDTGPKVTGNAALTLRNFCWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTSLQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480

```

Db      421 DCLLDQPSKPIITLPEDLPGTSYSLSQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV 480
Qy      481 QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR 540
Db      481 QTRHFPWADGTSCGEGKFLCKGACVERHNPKNKYRVDGSWAKWEPYGSCSRTCGGGVQLAR 540
Qy      541 RQCTNPPTANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
Db      541 RQCSNPPTANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
Qy      601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGTLTSPDSTSVQVQKCIKAGCD 660
Db      601 VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGTLTSPDSTSVQVQKCIKAGCD 660
Qy      661 GNLGSKKRFDKCGVCGDNKSCCKVFTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db      661 GNLGSKKRFDKCGVCGDNKSCCKRVFTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Qy      721 GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPPILEPLT 780
Db      721 GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPPILEPLT 780
Qy      781 VEVLVSGKMTPPRVYSFYLPKEPREDKSSHPKDPGRGSPVLHNSVLSLSNQVEQPDDRPP 840
Db      781 VEVLVSGKMTPPRVYSFYLPKEPREDKSTRPKDPRGSPVLHNSVLSLSNQVEQPDNRPP 840
Qy      841 ARWVAGSWGPCSASCGSGLQKRAVDGRSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900
Db      841 ARWVAGSWGPCSASCGSGLQKRAVDCRDSPGQAGASACDVRPLEKRACGEPCPTWELG 900
Qy      901 AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 950
Db      901 NWSPCSKSCGRGFQRRPLKCVGHGGRLRLARDQCDLRRKPQELDFCVLRPC 950

```

#### RESULT 4

##### Q4RYP9\_TETNG

```

ID   Q4RYP9_TETNG    PRELIMINARY;   PRT;   882 AA.
AC   Q4RYP9;
DT   19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT   19-JUL-2005, sequence version 1.
DT   07-FEB-2006, entry version 6.
DE   Chromosome 16 SCAF14974, whole genome shotgun sequence. (Fragment).
GN   ORFNames=GSTENG00026844001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15496914; DOI=10.1038/nature03025;
RA   Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.

```





Db	498	ENGGKYRCHGLRLKYSRCSLRPCPD--TGKSFRKQQCEEFNGLKLNTRLGSSVTWVPKYS	555
Qy	609	GVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIAKACDGNLGSKKR   :   :     : : :       :  : :     :   : :	668
Db	556	GVAPEDQCKLICRANGTGYFYVLAPKVVDGTPCSPTPSLCIQGRCIKAGCDGRNLNSRK	615
Qy	669	FDKCGVCGGDNKSCKKV'TGLFTKPMHGFNFVAIPAGASSIDIRQRGYKGLIGDDNYLAL      : : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	728
Db	616	FDKCGVCGGNNQCKKVSGRFAKPIRGYNFVV'TLPVGAANVDIRQRYGRGMSSDENYLAV	675
Qy	729	KNSQ GKYYLLNGHFVVS AVERDLVVGSLRLRYSGTGTAVESLQASRP ILEPLTVEVLSVGK  : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	788
Db	676	KDCRGHYLLNGGYVVSAAERDLLRVGGLMRYSGTSPA VET'LRAVAPLQEPLTVELLVSVG	735
Qy	789	MTPPRVRYSFYLP----KEPREDKSSH PKDP RGP SVLHNSVL SLSNQVEQPDDRPPA-RW      : : :   :   :   :   :   :   :   :   :   :   :   :   :   :	843
Db	736	MTPPRVRYT'FVYSVATGKEEEEEGRSH-----NYILEDEGRAEAGR W	777
Qy	844	VAGSWG PCSASC GSG LQ KRA VD CRGS AG QRT VPACDA AH RP V ETQACGE PCPTWELSAWS          :   :   :   :   :   :   :   :   :   :   :   :   :   :	903
Db	778	VAGGWQACSLTCGRGLQKRAVLQCDAEGR--AAD CDGAHRPLSERACGEPCPRWT'VG TWS	835
Qy	904	PCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELD F CVLRPC	950
Db	836	PCSTSCGRGFKRQRVC AVASGGQLPRERCAGLRKPQELDLCHLRTC	882

ATS15\_MOUSE

```

ID   ATSI15_MOUSE          STANDARD;          PRT;   562 AA.
AC   P59384;
DT   28-FEB-2003, integrated into UniProtKB/Swiss-Prot.
DT   28-FEB-2003, sequence version 1.
DT   07-MAR-2006, entry version 29.
DE   ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
DE   thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).
GN   Name=Adamts15;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=FVB/N; TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Bawa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   !- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   !- SUBCELLULAR LOCATION: Secreted protein; extracellular space;
CC   extracellular matrix (By similarity).
CC   !- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC   for a tight interaction with the extracellular matrix (By

```

AC P59384;

DT 28-FEB-2003, integrated into UniProtKB/Swiss-Prot.

DT 28-FEB-2003, sequence version 1.

DT 07-MAR-2006, entry version 29.

DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with  
DE thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).

GN Name=Adamts15;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidea; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC STRAIN=FVB/N; TISSUE=Kidney;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Huhli F.P., Jendry H., Meyer T., Moss S.I., Wang J., Hsieh F.

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang U., Hsiell F.,  
 B2 Diatchekas I., Marmaris K., Fayman J. J., Rubin G.M., Hong J.

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong D.,  
 B1 Stanletan M., Soares M.B., Bonalde M.E., Casavant M.I., Schoe

RA Brownstein M. J. Usdin T. B. Toshiyuki S. Carninci B. Brange C.

BA Baba S S Locuelli N A Peters G J Abramson B D Mullaby S

BA Bosak S. A. McEwan P. J. McKernan K. J. Malek J. A. Gunaratne P. H.

BA Richards S Worley K C Hale S Garcia A M Gay L J Hulyk S W

BA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., San

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial ana

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted protein

CC extracellular matrix (By similarity).

CC       -!- DOMAIN: The spacer domain and the ISF type-1 domains are  
CC       for a tight interaction with the extracellular matrix (By

```

CC      similarity).
CC      -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC      similarity).
CC      -!- SIMILARITY: Contains 1 disintegrin domain.
CC      -!- SIMILARITY: Contains 1 peptidase M12B domain.
CC      -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; BC043308; AAH43308.1; -; mRNA.
DR      HSSP; P07996; 1LSL.
DR      MEROPS; M12.025; -.
DR      Ensembl; ENSMUSG00000033453; Mus musculus.
DR      MGI; MGI:2449569; Adamts15.
DR      InterPro; IPR006586; ADAM_cysteine.
DR      InterPro; IPR010294; ADAM_spacer1.
DR      InterPro; IPR013273; ADAM_TS.
DR      InterPro; IPR001762; Disintegrin.
DR      InterPro; IPR001818; Pept_M10A_M12B.
DR      InterPro; IPR006025; Pept_M_Zn_BS.
DR      InterPro; IPR001590; Peptidase_M12B.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      Pfam; PF05986; ADAM_spacer1; 1.
DR      Pfam; PF00090; TSP_1; 3.
DR      PRINTS; PR01857; ADAMTSFAMILY.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00209; TSP1; 3.
DR      PROSITE; PS50215; ADAM_MEPRO; PARTIAL.
DR      PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR      PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR      PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
DR      PROSITE; PS50092; TSP1; 3.
DR      PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW      Extracellular matrix; Glycoprotein; Hydrolase; Metalloprotease;
KW      Protease; Repeat; Zinc.
FT      CHAIN          <1      562      ADAMTS-15.
FT                                     /FTid=PRO_0000078212.
FT      DOMAIN          <1      39      Peptidase M12B.
FT      DOMAIN          40     127      Disintegrin.
FT      DOMAIN          128     183      TSP type-1 1.
FT      DOMAIN          451     507      TSP type-1 2.
FT      DOMAIN          508     561      TSP type-1 3.
FT      REGION          <1      39      Metalloprotease.
FT      REGION          311     450      Spacer.
FT      COMPBias        184     312      Cys-rich.
FT      CARBOHYD         203     203      N-linked (GlcNAc . . .) (Potential).
FT      CARBOHYD         235     235      N-linked (GlcNAc . . .) (Potential).
FT      CARBOHYD         291     291      N-linked (GlcNAc . . .) (Potential).
FT      DISULFID         140     177      By similarity.
FT      DISULFID         144     182      By similarity.
FT      DISULFID         155     167      By similarity.
FT      DISULFID         463     502      By similarity.
FT      DISULFID         467     506      By similarity.
FT      DISULFID         478     490      By similarity.
FT      NON_TER          1        1
SQ      SEQUENCE        562 AA;  60964 MW;  48009AD337D44E4A CRC64;

```

```

Query Match          55.8%;  Score 2881;  DB 1;  Length 562;
Best Local Similarity 93.3%;  Pred. No. 1.9e-200;
Matches 517;  Conservative 16;  Mismatches 21;  Indels 0;  Gaps 0;

```

```

Qy      397  IDRANPWSACSAAIITDFLD SGHGDCLLDQPSKPISLPEDLP GASYTL SQQCELA FG VGS 456
          |||
Db      9   IDRANPWSACSAAIITDFLD SGHGDCLLDQPSKPI TL PEDLP GTSYSL SQQCELA FG VGS 68

Qy      457  KPCPYMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKLC LKGACVERHNLNKH RVD 516
          |||
Db      69  KPCPYMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKFLKGACVERHNPKNKY RVD 128

```

Qy	517	GSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASG	576
Db	129	GSWAKWEPYGCSTRTCGGGVQLARRQCSNPTPANGGKYCEGVRVKYRSCNLEPCPSSASG	188
Qy	577	KSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTYGYFVLAPKVV	636
Db	189	KSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTYGYFVLAPKVV	248
Qy	637	DGTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGY	696
Db	249	DGTLCTPDSTSVCVQGKCIKAGCDGNLGSKKKFDKCGVCGGDNKSCKRVTGLFTKPMHGY	308
Qy	697	NFVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGKYLNGHFVVSVERDLVVGKSL	756
Db	309	NFVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGKYLNGHFVVSVERDLVVGKSV	368
Qy	757	LRYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPRVRSFYLPKEPREDKSSHPKDPR	816
Db	369	LRYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPRVRSFYLPKEPREDKSTRPKDPR	428
Qy	817	GPSVLHNSVLSLSNQVEQPDDRPPARWAGSWGPCSASCGSGLQKRAVDCRGSAQRTVP	876
Db	429	GSPVLHNSVLSLSNQVEQPDNRPPARWAGSWGPCSVSCGSGLQKRAVDCRDSPGQQGAS	488
Qy	877	ACDAAHRPVETQACGEPCTWELSAWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLH	936
Db	489	ACDVDHRPLEKACGEPCTWELGNWSPCSKSCGRGFKRRPLKCVGHGGRLLARDQCDLR	548
Qy	937	RKPQELDFCVLRPC	950
Db	549	RKPQELDFCVLRPC	562



```

Db      569 VQGKCIKAGCDGKIGSDKKFDKCGICGGDNKGCKKVSGLFTKPVHGYNFVVMLPVGAANI 628
Qy      710 DIRQRGYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVESL 769
Db      629 DIRQRGYKGMTSDDNYLAVKNSDGRYLLNGNYIVSAGERDIIVKKSLLRYSGTTGLSETL 688
Qy      770 QASRPPILEPLTVEVLVSVGKMTPPRVYSFYLPKEPREDKS----SHPKDPGRGPSVLHNSV 825
Db      689 QSVKPLEEALTVEVLVSVGKMTPPRIRYSFYLSRQNKEDKNLKKKEAHDNSP-----NSV 741
Qy      826 LS-----LSNQVEQPDDRPPARWVAGSWGPCSASCGSLQKRAVDCRGSAQRTV 875
Db      742 LAHDGAKENGAKLLKSYNKENPAPGKWISAVWDKCSVTGNGLQRRSVQCLKPDGK--- 798
Qy      876 PA--CDAHRPVETQACGEPCPTWELSAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQC 933
Db      799 PATDCSSSQPSDTRVCGTPCEWLIGQWSPCSRTCGKGFRRPLHCKTQSGHSLPRDRC 858
Qy      934 NLHRKPQELDFCVLRPC 950
Db      859 TGLQKPQELDFCNLRSC 875

```

# RESULT 7

## Q2N1I7\_PIG

```

ID      Q2N1I7_PIG      PRELIMINARY;      PRT;      947 AA.
AC      Q2N1I7;
DT      07-FEB-2006, integrated into UniProtKB/TrEMBL.
DT      07-FEB-2006, sequence version 1.
DT      07-FEB-2006, entry version 1.
DE      ADAMTS1.
GN      Name=ADAMTS1;
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC      Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Yue K., Jiang S.;
RT      "Sus scrofa DNA for ADAMTS1.";
RL      Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; DQ177331; ABA54553.1; -; Genomic_DNA.
SQ      SEQUENCE 947 AA; 103009 MW; 86483441F9AF06D5 CRC64;

```

```

Query Match      48.5%; Score 2505; DB 2; Length 947;
Best Local Similarity 48.9%; Pred. No. 8e-173;
Matches 484; Conservative 160; Mismatches 246; Indels 100; Gaps 22;

```

```

Qy      1 MLLLGILT--LAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQIT 58
Db      17 LLLLTAAATLLVLQGAHGRPVEEDEELVLP-ALERDL-----AHGTAHLL--LD 62
Qy      59 AFQEDFYHLHTPDAQFLAPAFSTEHLG--VPLQGLTGSSDLRRCFYSGDVNAEPDSFA 115
Db      63 AFGSQLRLELQPDGRFLAPGFTLQTVGRRPGPNASHSDPAGDLAHCFSQTVNRDPSSAA 122
Qy      116 AVSLCGGLRGAFGYRGAEYVISPLPNAS----APAAQRNSQGA----HLLQRRGVPGGP 166
Db      123 ALSLCEGVRGAFYLGEEYFIQAPAAAAVGLAPAAAAGEEPLARPQHLLRRRRRRGGG- 181
Qy      167 SGDPTSRCGV-----ASGWNPAILRALDPYKPRRAGFGESRSRRRSRSG 208
Db      182 ----GAKCGVLDDETQLAKDAGSEGEDAAAQWPP-----QNWEPQRAG---QSTGTGSL 228
Qy      209 RAKRFVSIPIRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVVVK 268

```



RA Kuno K., Lizasa H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosomal mapping of the mouse  
 ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
 RL Genomics 46:466-471(1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=97150761; PubMed=8995297; DOI=10.1074/jbc.272.1.556;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
 RA Matsushima K.;  
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
 RT disintegrin family protein with thrombospondin motifs as an  
 RT inflammation associated gene.";  
 RL J. Biol. Chem. 272:556-562(1997).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Limb, and Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.  
 RX MEDLINE=99303657; PubMed=10373500; DOI=10.1074/jbc.274.26.18821;  
 RA Kuno K., Terashima Y., Matsushima K.;  
 RT "ADAMTS-1 is an active metalloproteinase associated with the  
 RT extracellular matrix.";  
 RL J. Biol. Chem. 274:18821-18826(1999).  
 RN [5]  
 RP FUNCTION.  
 RX MEDLINE=20389568; PubMed=10930576; DOI=10.1016/S0014-5793(00)01854-8;  
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M., Ohno H.,  
 RA Matsushima K.;  
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
 RL FEBS Lett. 478:241-245(2000).  
 RN [6]  
 RP FUNCTION, AND INDUCTION.  
 RX MEDLINE=20243757; PubMed=10781075; DOI=10.1073/pnas.080073497;  
 RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
 RA Richards J.S.;  
 RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and  
 RT cathepsin L proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
 CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. Has angiogenic inhibitor activity (By  
 CC similarity). Active metalloprotease, which may be associated with  
 CC various inflammatory processes as well as development of cancer  
 CC cachexia. May play a critical role in follicular rupture (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692  
 CC site, within the chondroitin sulfate attachment domain.  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted protein; extracellular space;  
 CC extracellular matrix.  
 CC -!- INDUCTION: Induced in vitro in colon adenocarcinoma cells by



CC interleukin-1, or in vivo in kidney and heart by  
 CC lipopolysaccharide. Also induced by LH stimulation in granulosa  
 CC cells of preovulatory follicles.  
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase.  
 CC -!- SIMILARITY: Contains 1 disintegrin domain.  
 CC -!- SIMILARITY: Contains 1 peptidase M12B domain.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 7.

CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AB001735; BAA24501.1; ALT\_INIT; Genomic\_DNA.  
 DR EMBL; D67076; BAA11088.1; ALT\_FRAME; mRNA.  
 DR EMBL; BC040382; AAH40382.1; -; mRNA.  
 DR EMBL; BC050834; AAH50834.1; -; mRNA.  
 DR HSSP; P07996; 1LSL.  
 DR MEROPS; M12.222; -.  
 DR Ensembl; ENSMUSG00000022893; Mus musculus.  
 DR MGI; MGI:109249; Adamts1.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01858; ADAMTS1.  
 DR PRINTS; PR01857; ADAMTSFAMILY.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
 KW Metal-binding; Metalloprotease; Protease; Repeat; Signal; Zinc;  
 KW Zymogen.  
 FT SIGNAL 1 48 Potential.  
 FT PROPEP 49 253  
 FT CHAIN 254 968 /FTId=PRO\_0000029152.  
 FT ADAMTS-1.  
 FT /FTId=PRO\_0000029153.  
 FT DOMAIN 259 468 Peptidase M12B.  
 FT DOMAIN 477 559 Disintegrin.  
 FT DOMAIN 560 615 TSP type-1 1.  
 FT DOMAIN 855 911 TSP type-1 2.  
 FT DOMAIN 912 968 TSP type-1 3.  
 FT REGION 726 850 Spacer.  
 FT COMPBias 195 199 Poly-Arg.  
 FT COMPBias 618 725 Cys-rich.  
 FT ACT\_SITE 403 403  
 FT METAL 402 402 Zinc (catalytic) (By similarity).  
 FT METAL 406 406 Zinc (catalytic) (By similarity).  
 FT METAL 412 412 Zinc (catalytic) (By similarity).  
 FT SITE 206 206 Cysteine switch (Potential).  
 FT CARBOHYD 548 548 N-linked (GlcNAc . . .) (Potential).  
 FT CARBOHYD 721 721 N-linked (GlcNAc . . .) (Potential).  
 FT CARBOHYD 765 765 N-linked (GlcNAc . . .) (Potential).

FT	CARBOHYD	783	783	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	946	946	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	380	463	By similarity.
FT	DISULFID	418	447	By similarity.
FT	DISULFID	572	609	By similarity.
FT	DISULFID	576	614	By similarity.
FT	DISULFID	587	599	By similarity.
FT	MUTAGEN	403	403	E->Q: Loss of activity.
FT	CONFLICT	335	335	N -> S (in Ref. 2).
FT	CONFLICT	425	425	T -> S (in Ref. 2).
SQ	SEQUENCE	968 AA;	105842 MW;	42EBDA55499FB6C1 CRC64;

Query Match 48.1%; Score 2482.5; DB 1; Length 968;  
 Best Local Similarity 48.0%; Pred. No. 3.5e-171;  
 Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

Qy	1	MLLLGILTLAGRTAGG--FEPEREVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI	57
		:    : :         : : :    :         : :	
Db	37	LLLLASITMLLCARGAHGRPTEDEELVLP-SLE-----RAPGHDSTTTRL--RL	83
Qy	58	TAFQEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGs-----SDLRRCFYSGDVNA	109
		: :            : : :   :	
Db	84	DAFGQQLHLKLQPD SGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHC FYSGTVNG	138
Qy	110	EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS----APAAQRNSQGA----HLLQRRG	161
		:                    :         :	
Db	139	DPGSAAALSLCEGVRGAFYLGQEEFFIQPAPGVATERLAPAVPEEESARPQFHILRRR-	197
Qy	162	VPGGPSGDPTSRCGVASGWNPAILRALDPYKPRRAGFGESRSRR-----206	
		: :	
Db	198	----RRGSGGAKCGVMD-----DETLP TSDSRPESQNT RNQWPVRDPTPDAGKP	243
Qy	207	----SGRAKR FVSIPRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNPI	262
Db	244	SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHLYLLTFSVAARFYKHP SIRSNI	303
Qy	263	NIVVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVDKHP EYWDTAILFTRQDLC	322
		: :            : : :	
Db	304	SLVVVKILVIYEEQKGEVTSNAALTLRNFCNWQKHNSPSDRDPEHYDTAILFTRQDLC	363
Qy	323	GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFG	382
		:	
Db	364	GSHTCDTLGMADVGTVC DPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCASLNG	423
Qy	383	KLRANHHMMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPI SLPEDLP GASY	442
		: :  :  :   :   :         : :	
Db	424	VTGDSHLMASMLSSLDHSQPWSPCSAYMVT SFLDNHGHECLMDKPNPIKLP S DLP GTLY	483
Qy	443	TLSQQCELA FGVGSKPCP-YMQYCTKLWCTGKAKGQMV CQTRHFPWADGTSCGEGKCLK	501
		: :  :  :                     : :	
Db	484	DANRQCQFTFGEESKHCPDAASTCTTLWCTGTSGGLLV CQTKHFPWADGTSCGEGKWCVS	543
Qy	502	GACVERHNLNKH--RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV	558
		: :	
Db	544	GKCVNK TDM-KHFATPVHGSWGPWGPWGDCSRTCGGGVQYTMRECDNVPVKNGGKYCEGK	602
Qy	559	RVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDCKKL	618
		:                 : :                 : :	
Db	603	RVRYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL	661
Qy	619	ICRANGTG YFYVLAPKVVDGTL CSPDSTSV CVQGKCIKAGCDGNLGSKKRFDKCGVCGGD	678
Db	662	TCEAKGIGYFFVLQPKVVDGTPCSPDSTSV CVQGQCVKAGCDRIIDSKKKFDKCGVCGGN	721
Qy	679	NKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLN	738
		:     : :  :             : :   :   : :	
Db	722	GSTCKKMSGIVTSTRPGYHDI VTI PAGATNIEVKHRNQRGSRNNGSFLAIRAADGTYILN	781
Qy	739	GHFVVS AVERDLVVKGSLRLYS GTGTAVESLQASRPILEPLTVEVLSVGKMTPPRVRYSF	798
		:   :   :                 : :   :   :	
Db	782	GNFTLSTLEQDLTYKGTVLRYSGSSAALERIRSFSPLKEPLTIQVLMVGHALRPKIKFTY	841

Qy	799	YLPKEPREDKSSHPKDPGRGPSVLHNSVLSLSNQVEQPDDRRPPARWVAGSWGPCSASCGSG	858
		::   :   : :           :	
Db	842	FMKKKTES-----FNAIPTFS-----EWVIEEWGECSTCTCGSG	874
Qy	859	LQKRAVDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCGRGFQR	915
		:           :   :   :   :         :   :   :	
Db	875	WQRRVVQCRDINGH--PASECAKEVKPASTRPCADLPCHWQVGDWSPCKTCKGKYKK	931
Qy	916	RLSKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC	950
		:           :   :   :   :	
Db	932	RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC	967

03U0I4 MOUSE

```

ID Q3U0I4_MOUSE PRELIMINARY; PRT; 968 AA.
AC Q3U0I4;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Activated spleen cDNA, RIKEN full-length enriched library,
DE clone:F830102H03 product:a disintegrin-like and metalloprotease
DE (reprolysin type) with thrombospondin type 1 motif, 1, full insert
DE sequence.
GN Name=Adamts1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminiecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovskiy E., Zhu S., Zimmer A., Hide W., Bult C.,

```

RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";   
 RL Science 309:1559-1563(2005).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 RG (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";   
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";   
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD; TISSUE=Activated spleen;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,  
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AK156830; BAE33869.1; -; mRNA.  
 DR MGI; MGI:109249; Adamts1.  
 DR GO; GO:0031410; C:cytoplasmic vesicle; IDA.  
 DR GO; GO:0031012; C:extracellular matrix; IDA.  
 DR GO; GO:0008201; F:heparin binding; IDA.  
 DR GO; GO:0001822; P:kidney development; IMP.  
 DR GO; GO:0001542; P:ovulation (sensu Mammalia); IMP.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Integrin; Metalloprotease; Protease.  
 SQ SEQUENCE 968 AA; 105826 MW; 7FB0B6554984DACA CRC64;

Query Match 48.1%; Score 2482.5; DB 2; Length 968;

Best Local Similarity 48.0%; Pred. No. 3.5e-171;  
Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

```
Qy      1 MLLLGILTAFAGRTAGG--FEPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
      :||| :|: | | | : | :| | : | : | : | : | : | : | : | : | : | : |
Db      37 LLLLASITMLLCARGAHRPTEEDEELVLP-SLE-----RAPGHDSITTRL--RL 83

Qy      58 TAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQGLTGGG-----SDLRRCFYSGDVNA 109
      || : :| | | : || | : : : | : || | | | | | | |
Db      84 DAFGQQLHLKLQPDGFLAPGFTLQTV----GRSPGSEAQHLDPTGDLAHCIFYSGTVNG 138

Qy     110 EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQRRG 161
      :| | | :|| | :||| | :| | : | | : | | | | | :| :| |
Db     139 DPGSAAALSLCEGVRGAFYLGQEEFFIQPAPGVATERLAPAVPEEESARPQFHILRRR- 197

Qy     162 VPGGPGSDPTSRGCVASGWNPAILRALDPYKPRRAGFGESRSRR----- 206
      | :||| | | | | | | :| : |
Db     198 ---RRGSGGAKCGVMD-----DETLPSTDSRPESQNRNQWVPRDPTPDAGKP 243

Qy     207 ---SGRAKRFVSIPIRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNPI 262
      | | |||| | |||| :||| : | | : | | | | : | | | | | | |
Db     244 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHYLLTLFSVAARFYKHPISIRNSI 303

Qy     263 NIVVVVLLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVSDDHPEYWDTAILFTRQDLC 322
      :||| :| : | : | | | | | | | | | | | | | | | | | | |
Db     304 SLVVVVKILVIYEEQKGPVTSNAALTLRNFCNWQKQHNPSDRDPEHYDTAILFTRQDLC 363

Qy     323 GATTCDTLGMADVGTMDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCCEEVFG 382
      | : ||||| | | | | | | | | | | | | | | | | | | : | : |
Db     364 GSHTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCASLNG 423

Qy     383 KLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442
      :| :| : | : | | | | :| | | :| :| :| :| :| | | | | |
Db     424 VTGDSHLMASMLSSLDHSQPWSPCSAYMVSFLDNHGHGECMLDKPQNPIKPSDLPGTLY 483

Qy     443 TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKLCLK 501
      :|| : | | | | | | | | | | | | | | | | | | | | :
Db     484 DANRQCQFTFGEEKHCPDAASTCTTLWCTGTSGGLLVCQTKHFPWADGTSCGEGKWCVS 543

Qy     502 GACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV 558
      | | | : : | | | | | | | | | | | | | | | | | | |
Db     544 GKCVNKTDM-KHFATPVHGSWGPWGDSCSRTCGGGVQYTMRECDNPVPKNGGKYCEGK 602

Qy     559 RVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKL 618
      || :||| : | : :||| | | : : | | | | :| :| :| :|
Db     603 RVRYSRNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 661

Qy     619 ICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGGKCIKAGCDGNLGSKKRFDKCGVCGGD 678
      | | | ||| :| | | | | | | | | | | | | | : | | :| | | | :
Db     662 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGGQCVKAGCDRIIDSKKKFDKCGVCGGN 721

Qy     679 NKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGKYLLN 738
      :||| :| : | : | | | :| :| : | : | : | : | : | : |
Db     722 GSTCKKMSGIIVTSTRPGYHDIVTIPAGATNIEVKHRNQGRSRNNGSFLAIRAADGTIILN 781

Qy     739 GHFVVSAYERDLVVKGSLRLRYSGTGTAVESLQASRPILPTVEVLVSGKMTPPRVRYSF 798
      | : | :| :| | :| :||| : | :| :| :| :| | | :| :| :|
Db     782 GNFTLSTLEQDLTYKGTVLRYSGSSAALERIRSFSPLEPLTIQVLMVGHAPRPKIKFTY 841

Qy     799 YLPKEPREDKSSHKDPGSPVLHNSVLSLNSQVEQPDDRPAPRWVAGSWGSPCSASCGSG 858
      : : | : | : | : | : | : | : | : | : | : | : |
Db     842 FMKKKTES-----FNAIPTFS-----EWVIEEWGECSTCGSG 874

Qy     859 LQKRAVDRCGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915
      | : | | | | | : | : | : | : | : | : | : | : | : | :
Db     875 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCSKTCGKGYKK 931

Qy     916 RSLKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
      | :||| | | :| : | : | :| : | | |
Db     932 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 967
```

RESULT 10

AT51\_HUMAN

ID AT51\_HUMAN STANDARD; PRT; 967 AA.  
AC Q9UHI8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;  
DT 30-MAY-2000, integrated into UniProtKB/Swiss-Prot.  
DT 01-DEC-2000, sequence version 3.  
DT 07-MAR-2006, entry version 60.  
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).  
GN Name=ADAMTS1; Synonyms=KIAA1346, METH1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;  
RT "Cloning, characterization and mapping on human chromosome 21 of the  
RT orthologue of murine Adamts-1.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.  
RC TISSUE=Heart;  
RX MEDLINE=99367466; PubMed=10438512; DOI=10.1074/jbc.274.33.23349;  
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
RA Lombardo M., Iruela-Arispe M.L.;  
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
RT family of proteins with angio-inhibitory activity.";  
RL J. Biol. Chem. 274:23349-23357(1999).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RC TISSUE=Endothelial cell;  
RX MEDLINE=20247184; PubMed=10785405;  
RA Glienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,  
RA Rosenthal A., Thierach K.H.;  
RT "Differential gene expression by endothelial cells in distinct  
RT angiogenic states.";  
RL Eur. J. Biochem. 267:2820-2830(2000).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Brain;  
RX MEDLINE=20181126; PubMed=10718198; DOI=10.1093/dnares/7.1.65;  
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVI.  
RT The complete sequences of 150 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:65-73(2000).  
RN [5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;  
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,  
RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,  
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,  
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
RA Minoshima S., Shimizu N., Nordtsiek G., Hornischer K., Brandt P.,  
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,  
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,  
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
RA Lehrach H., Reinhardt R., Yaspo M.-L.;  
RT "The DNA sequence of human chromosome 21.";  
RL Nature 405:311-319(2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 418-967.  
RC TISSUE=Melanoma;  
RG The German cDNA consortium;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover (By similarity). Has angiogenic inhibitor  
 CC activity. Active metalloprotease, which may be associated with  
 CC various inflammatory processes as well as development of cancer  
 CC cachexia. May play a critical role in follicular rupture.  
 CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1938-Glu-|-Leu-1939  
 CC site, within the chondroitin sulfate attachment domain.  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted protein; extracellular space;  
 CC extracellular matrix (By similarity).  
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -!- SIMILARITY: Contains 1 disintegrin domain.  
 CC -!- SIMILARITY: Contains 1 peptidase M12B domain.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.

-----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AF170084; AAF15317.1; -; mRNA.  
 DR EMBL; AF060152; AAD48080.1; ALT\_INIT; mRNA.  
 DR EMBL; AF207664; AAF23772.1; -; mRNA.  
 DR EMBL; AB037767; BAA92584.1; ALT\_INIT; mRNA.  
 DR EMBL; AP001697; BAA95502.1; -; Genomic\_DNA.  
 DR EMBL; AL162080; CAB82413.1; -; mRNA.  
 DR PIR; T47158; T47158.  
 DR HSSP; P07996; 1LSL.  
 DR MEROPS; M12.222; -.  
 DR Ensembl; ENSG00000154734; Homo sapiens.  
 DR H-InvDB; HIX0016042; -.  
 DR HGNC; HGNC:217; ADAMTS1.  
 DR MIM; 605174; gene.  
 DR GO; GO:0008237; F:metallopeptidase activity; TAS.  
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP1; 3.  
 DR PRINTS; PR01858; ADAMTS1.  
 DR PRINTS; PR01857; ADAMTSFAMILY.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.

KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
 KW Metal-binding; Metalloprotease; Protease; Repeat; Signal; Zinc;  
 KW Zymogen.

FT	SIGNAL	1	49	Potential.
FT	PROPEP	50	252	By similarity.
FT				/FTId=PRO_0000029150.
FT	CHAIN	253	967	ADAMTS-1.
FT				/FTId=PRO_0000029151.
FT	DOMAIN	258	467	Peptidase M12B.
FT	DOMAIN	476	559	Disintegrin.



FT	DOMAIN	559	614	TSP type-1 1.
FT	DOMAIN	854	905	TSP type-1 2.
FT	DOMAIN	908	967	TSP type-1 3.
FT	REGION	725	849	Spacer.
FT	COMPBias	617	724	Cys-rich.
FT	COMPBias	843	846	Poly-Lys.
FT	ACT_SITE	402	402	By similarity.
FT	METAL	401	401	Zinc (catalytic) (By similarity).
FT	METAL	405	405	Zinc (catalytic) (By similarity).
FT	METAL	411	411	Zinc (catalytic) (By similarity).
FT	SITE	198	198	Cysteine switch (Potential).
FT	CARBOHYD	547	547	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	720	720	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	764	764	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	379	462	By similarity.
FT	DISULFID	417	446	By similarity.
FT	DISULFID	571	608	By similarity.
FT	DISULFID	575	613	By similarity.
FT	DISULFID	586	598	By similarity.
FT	CONFLICT	227	227	P -> A (in Ref. 4 and 5).
FT	CONFLICT	468	468	Q -> H (in Ref. 1).
FT	CONFLICT	561	561	S -> N (in Ref. 1).
SQ	SEQUENCE	967 AA;	105384 MW;	C189389324741ED1 CRC64;

Query Match 48.0%; Score 2480.5; DB 1; Length 967;  
 Best Local Similarity 48.9%; Pred. No. 5e-171;  
 Matches 485; Conservative 154; Mismatches 251; Indels 101; Gaps 24;

Qy	1	MLLLGIILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	36	LLLLAAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTTRLRLHAF	81
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF	114
Db	82	DQQLDLELRPDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA	136
Qy	115	AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA----HLLQ--RRGVPGG	165
Db	137	AALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLRLRRNRQGDVGG	196
Qy	166	PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFGESRSRRRS	207
Db	197	TCGVVDDEPRPTGKAETEDDEGTEGEDEGPQWS-----PQDPALQGVGQP-TGTGS	247
Qy	208	GRAKRFSVIPRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVVV	267
Db	248	IRKKRFVSSHRYVETMLVADQSMAEFHGSGLKHLLTLFSVAARLYKHPSIRNSVSLVVV	307
Qy	268	KVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVS DKHPEYWDTAILFTRQDLCGATTC	327
Db	308	KILVIHDEQKGPEVTSNAALTLRNFCAWQKHNPSPDRDAEHYDTAILFTRQDLCSQTC	367
Qy	328	DTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVC EEVFGKLRAN	387
Db	368	DTLGMADVGTVCDPSPRSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKQCASLNGVNQDS	427
Qy	388	HMSPTLIQIDRANPWSACSAAIITDFLD SGHGDCLLDQPSKPISLPEDLPGASYTLSQQ	447
Db	428	HMMASMLSNLDHSQPWSPCSAYMITSFLDNHGGECLMDKPQNPIQLPGDLPGTSYDANRQ	487
Qy	448	CELAFGVGSKPCP-YMQYCTKLWCTGKAKGQMVQC TRHFPWADGTSCGEGKLCLKGACVE	506
Db	488	CQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWCINGKCVN	547
Qy	507	RHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNP TPANGGKYCEGVRVKYR	563
Db	548	KTD-RKHFDTPFHGSWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYR	606
Qy	564	SCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDCKKLCICRAN	623
Db	607	SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKKLCICAK	665

Qy 624 GTGYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCK 683  
 Db 666 GIGYFFVLQPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 725

Qy 684 KVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLNGHFVV 743  
 Db 726 KISGSVTSAPGYHDIITIPATGATNIEVKQRNQGRSRNGSFLAIKAADGTIILNGDYTL 785

Qy 744 SAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPTVEVLSVGKMTPPRVRYSFYLPKE 803  
 Db 786 STLEQDIMYKGVVLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKIKYTYFVKKK 845

Qy 804 PREDKSSHPKDPGRPSVLHNSVLSLNSQVEQDDPPARWVAGSWGPCSASCGSGLQKRA 863  
 Db 846 ---KES-----FNAIPTFS-----AWVIEEWGECSKSCELGWQRR 878

Qy 864 VDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQRRSLKC 920  
 Db 879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSKTCGKGYKKRSLKC 935

Qy 921 VGHGGRLRLARDQCNLHRKPQE-LDFCVLRPC 950  
 Db 936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966

# RESULT 11

## Q5HYL0\_HUMAN

ID Q5HYL0\_HUMAN PRELIMINARY; PRT; 967 AA.  
 AC Q5HYL0;  
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.  
 DT 15-FEB-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 4.  
 DE Hypothetical protein DKFZp686E01144.  
 GN Name=DKFZp686E01144;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Colon endothel;  
 RG The German cDNA Consortium;  
 RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,  
 RA Fobo G., Han M., Wiemann S.;  
 RL Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.

CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; BX647388; CAI46043.1; -; mRNA.  
 DR Ensembl; ENSG00000154734; Homo sapiens.  
 DR GO; GO:0031012; C:extracellular matrix; IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis; IEA.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00608; ACR; 1.  
DR SMART; SM00209; TSP1; 3.  
DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
DR PROSITE; PS50092; TSP1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
KW Hypothetical protein.  
SQ SEQUENCE 967 AA; 105354 MW; 24685FFB095A9497 CRC64;

Query Match 48.0%; Score 2480.5; DB 2; Length 967;  
Best Local Similarity 48.9%; Pred. No. 5e-171;  
Matches 485; Conservative 154; Mismatches 251; Indels 101; Gaps 24;

Qy 1 MLLLGILTLAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
: ||| : | : : ||| : | | : : ||  
Db 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTARLRLHAF 81

Qy 61 QEDFYLHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF 114  
: | | | | : ||| : : : | | : || |||| | : | |  
Db 82 DQQLDLELRPDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 136

Qy 115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQ--RRGVPGG 165  
||| | | | | : | | | | | | | | | | | : | : | |  
Db 137 AALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLLRRNRQGDVGG 196

Qy 166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFGESRSRRRS 207  
| | : | | : | | : | : |  
Db 197 TCGVVDDEPRPTGKAETEDEDEGTEGEDEGPQWS-----PQDPALQGVGQP-TGTGS 247

Qy 208 GRAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVV 267  
| | | | : | | : | | : | | : | | | | : | | | | | : : | |  
Db 248 IRKKRFVSSHRYVETMLVADQSMAEFHGSGLKHYLLTLFSAARLYKHPSIRNSVSLVV 307

Qy 268 KVLRLDRDSGPKVTGNAALTLRNFCAWQKKLNKVS DKHPEYWDTAILFTRQDLCGATTC 327  
| : : | : | : | | | | | | : | : | : | | | | | | : |  
Db 308 KILVIHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQTC 367

Qy 328 DTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVC EEVFGKLRAN 387  
| | | | | : | | | | | | | | | | | : | : | :  
Db 368 DTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNQDS 427

Qy 388 HMMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTLSQQ 447  
| | : | : | : | | : | | | : | : | : | : | | | | : |  
Db 428 HMMASMLSNLDHSQPWSPCSAYMITSFLDNHGECLMDKPQNPIQLPGDLPGTSYDANRQ 487

Qy 448 CELAFGVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHPWADGTSCGEGKCLKGACVE 506  
| : | | | | : | | | : | : | : | | | | | | | : | |  
Db 488 CQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHPWADGTSCGEGKWCINGKCVN 547

Qy 507 RHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYR 563  
: : | | | | : | | | | | : | : | | | | | : | |  
Db 548 KTD-RKHFDTPFHGSWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVYR 606

Qy 564 SCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKLICRAN 623  
| | | | | : | : | | | | : : | | : | | : | : | : |  
Db 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAK 665

Qy 624 GTGYFYVLAPKVVDGTLCS PDSTSVCVQGGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCK 683  
| | | : | | | | | | | | : | | | : | | : | | | : | |  
Db 666 GIGYFFVLQPKVVDGTPCSPDSTSVCVQGGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 725

Qy 684 KVTGLFTKPMHGYNFVVAIPAGASSIDIRGRYKGLIGDDNYLALKNSQGKYLLNGHFVV 743  
| : | | | : : | | | : : | : | : : | : | : | : | :  
Db 726 KISGSVTSAPKGYHDIITPTGATNIEVKQRNQRGSRNNGSFLAIKAADGTYILNGDYTL 785

Qy 744 SAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPTVEVLVSGKMTPPRVRYSFYLPKE 803  
| : | : : | : | | : | : : | : | | : | : | : : | :  
Db 786 STLEQDIMYKGVVLRYSGSSAALERIRSFSPLKEPLTIQVLTVGNALRPKIKYTYFVKKK 845

Qy 804 PREDKSSHPKDPGSPVLHNSVLSLSNQVEQPDDRPPARVWAGSWGPCSASCGSGLQKRA 863  
| | : : | | | | | | : |

Db 846 ----KES-----FNAIPTFS-----AWVIEEWGECSKSCSELGWQRRRL 878

QY 864 VDCRGSAQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQRRSLKC 920  
 ||| || || : | : | : ||| ||| |||:|:|:|:|:|

Db 879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSKTCGKGKKRSLKC 935

QY 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950  
 : | | :|: : | : ||| : ||

Db 936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966

# RESULT 12

## Q5U261\_XENLA

ID Q5U261\_XENLA PRELIMINARY; PRT; 928 AA.

AC Q5U261;

DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.

DT 07-DEC-2004, sequence version 1.

DT 07-FEB-2006, entry version 10.

DE LOC495679 protein.

GN Name=LOC495679;

OS *Xenopus laevis* (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;

OC Xenopodinae; *Xenopus*;

OX NCBI\_TaxID=8355;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Embryo;

RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;

RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RT "Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus*

RT initiative.";

RL Dev. Dyn. 225:384-391(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Embryo;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Embryo;

RA Klein S., Gerhard D.S.;

RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; BC086266; AAH86266.1; -; mRNA.

DR GO; GO:0031012; C:extracellular matrix; IEA.

DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis; IEA.



Qy 711 IRQRGYKGLIGDDNYLALKNSQGKYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQ 770  
 Db 703 VKQRNNRGRSRHDGSLAIIKAADGTYLNGDYTLSTLEQDITHNGNVLRYSGSSASLERIR 762

Qy 771 ASRPILEPLTVEVLSVGKMTPPRVRYSFYLPKEPREDKSSHPKDPGRGPSVLHNSVLSLSN 830  
 Db 763 SFSPLKEPITIQVLTVGDSHRLKIKYVYFVKKTGQPEKPNKKKE--SFNAIRETILS--- 817

Qy 831 QVEQPDDRPPARWVAGSWGSPCSASCGSLQKRAVDCRGSAQRTVPACDAAH--RPVETQ 888  
 Db 818 -----EWVIEEWGECSKTCGLGWQRRKVECKDINGQ---PSMDCANELKPDDIR 863

Qy 889 ACGE-PCPTWELSAWSPCSKSCGRGFQRRSLKCVGHGRLRLARDQCNLHRKPQEL-DFCV 946  
 Db 864 PCADTPCPQWQLGDWSSCSKTCGKGFKKRLKCVSYDGVNMPQENCDSLKKPKHLIDFCT 923

Qy 947 LRPC 950  
 Db 924 LANC 927

# RESULT 13

## Q8NE26\_HUMAN

ID Q8NE26\_HUMAN PRELIMINARY; PRT; 967 AA.  
 AC Q8NE26;  
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.  
 DT 01-OCT-2002, sequence version 1.  
 DT 07-FEB-2006, entry version 18.  
 DE ADAM metalloproteinase with thrombospondin type 1 motif, 1,  
 DE preproprotein.  
 GN Name=ADAMTS1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RG NIH MGC Project;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC036515; AAH36515.1; -; mRNA.

DR HSSP; P07996; 1LSL.  
 DR Ensembl; ENSG00000154734; Homo sapiens.  
 DR GO; GO:0031012; C:extracellular matrix; IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis; IEA.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 SQ SEQUENCE 967 AA; 105388 MW; FF1D399674201C3D CRC64;

Query Match 47.9%; Score 2472.5; DB 2; Length 967;  
 Best Local Similarity 48.8%; Pred. No. 1.9e-170;  
 Matches 484; Conservative 154; Mismatches 252; Indels 101; Gaps 24;

Qy	1	MLLLGILTLAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
		:       :   :      :       :	
Db	36	LLLLAAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTRLRLHAF	81
Qy	61	QEDFYHLHTPDQAFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF	114
		:       :       : :                 :             :	
Db	82	DQQLDLELRPDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA	136
Qy	115	AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA----HLLQ--RRGVPGG	165
		:       :                                  :   :	
Db	137	AALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLLRRNRQGDVGG	196
Qy	166	PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFGESRSRRRS	207
		:       :       :       :	
Db	197	TCGVVDDEPRPTGKAETEDEGTEGEDEGPQWS-----PQDPALQGVGQP-TGTGS	247
Qy	208	GRAKRFSVIPRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVV	267
		:    :   :    :       :       :	
Db	248	IRKKRFVSSHRYVETMLVADQSMAEFHGSGLKHYLLTLFSVAARLYKHPSIRNSVSLVVV	307
Qy	268	KVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVS DKHPEYWDTAILFTRQDLCGATTC	327
		:    :     :              :    :   :             :	
Db	308	KILVIHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQTC	367
Qy	328	DTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVC EEVFGKLRAN	387
		:               :                 :   :   :	
Db	368	DTLGMADVGTVC DPSRSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKQCASLNGVNQDS	427
Qy	388	HMSPTLIQIDRANPWSACSAAIITDFLD SGHGDCLLDQPSKPI SLPEDLP GASYTLSQQ	447
		:   :   :      :       :    :                :	
Db	428	HMMASMLSNLDHSQPWSPCSAYMITSFLDNHGGECLMDKPQNPIQLPGDLPGTSYDANRQ	487
Qy	448	CELAFGVGSGKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKLC LKGACVE	506
		:            :      :   :     :               :	
Db	488	CQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWCINGKCVN	547
Qy	507	RHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNP TPANGGKYCEGVRVKYR	563
		: :           :           :                  :	
Db	548	KTD-RKHFDTPFHGSGMWPWGDCSRTCGGGVQYTMRECDNPV PKNGGKYCEGKRVRYR	606

```

Qy      564 SCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKLICRAN 623
Db      607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAK 665

Qy      624 GTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSKC 683
Db      666 GIGYFFVLQPKVVDGTPCSTDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 725

Qy      684 KVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLNGHFVV 743
Db      726 KISGSVTSAPKPYHDIITPTGATNIEVKQRNQRGSRNNGSFLAIIAADGTIILNGDYTL 785

Qy      744 SAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPRVRSFYLPKE 803
Db      786 STLEQDIMYKGVVLRYSGSSAALERIRSFSLKEPLTIQVLTVGNALRPKIKYTYFVKKK 845

Qy      804 PREDKSSHPKDPGRGPSVLHNSVLSLSNQVEQPDDRPPARWVAGSWGPCSASCGSGLQKRA 863
Db      846 ---KES-----FNAIPTFS-----AWVIEEWGECSSKSELGWQRRRL 878

Qy      864 VDCRGSAQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQRRSLKC 920
Db      879 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSKTCGKGYKKRSLKC 935

Qy      921 VGHGGRLRLARDQCNLHRKPQE-LDFCVLRPC 950
Db      936 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 966

```

#### RESULT 14

##### ATSl\_RAT

```

ID   ATSl_RAT          STANDARD;          PRT;   967 AA.
AC   Q9WUQ1; Q9ERI1;
DT   01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT   01-NOV-1999, sequence version 1.
DT   07-MAR-2006, entry version 57.
DE   ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN   Name=Adamts1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RA   Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA   Little S.P.;
RT   "Induction of a disintegrin and metalloprotease with the
RT   thrombospondin type I motif (ADAMTS).";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [MRNA] OF 18-967.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=20304099; PubMed=10847486;
RX   DOI=10.1034/j.1600-0676.2000.020002165.x;
RA   Diamantis I., Luethi M., Hoesli M., Reichen J.;
RT   "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT   endothelial cells in cirrhotic rats.";
RL   Liver 20:165-172(2000).
CC   -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC   involved in its turnover. Has angiogenic inhibitor activity (By
CC   similarity). Active metalloprotease, which may be associated with
CC   various inflammatory processes as well as development of cancer
CC   cachexia. May play a critical role in follicular rupture (By
CC   similarity).
CC   -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684
CC   site, within the chondroitin sulfate attachment domain.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

```



CC -!- SUBCELLULAR LOCATION: Secreted protein; extracellular space;  
 CC extracellular matrix (By similarity).  
 CC -!- INDUCTION: Down-regulated in endothelial cells derived from  
 CC cirrhotic liver.  
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -!- SIMILARITY: Contains 1 disintegrin domain.  
 CC -!- SIMILARITY: Contains 1 peptidase M12B domain.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.

CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AF149118; AAD34012.1; -; mRNA.  
 DR EMBL; AF304446; AAG29823.1; -; mRNA.  
 DR HSSP; P07996; 1LSL.  
 DR MEROPS; M12.222; -.  
 DR Ensembl; ENSRNOG00000001607; Rattus norvegicus.  
 DR RGD; 621241; Adamts1.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01858; ADAMTS1.  
 DR PRINTS; PR01857; ADAMTSFAMILY.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;  
 KW Metal-binding; Metalloprotease; Protease; Repeat; Signal; Zinc;  
 KW Zymogen.

FT	SIGNAL	1	54	Potential.
FT	PROPEP	55	252	By similarity.
FT				/FTId=PRO_0000029154.
FT	CHAIN	253	967	ADAMTS-1.
FT				/FTId=PRO_0000029155.
FT	DOMAIN	258	467	Peptidase M12B.
FT	DOMAIN	476	558	Disintegrin.
FT	DOMAIN	559	614	TSP type-1 1.
FT	DOMAIN	854	910	TSP type-1 2.
FT	DOMAIN	911	967	TSP type-1 3.
FT	REGION	725	857	Spacer.
FT	COMBIAS	194	198	Poly-Arg.
FT	COMBIAS	616	724	Cys-rich.
FT	ACT_SITE	402	402	By similarity.
FT	METAL	401	401	Zinc (catalytic) (By similarity).
FT	METAL	405	405	Zinc (catalytic) (By similarity).
FT	METAL	411	411	Zinc (catalytic) (By similarity).
FT	SITE	205	205	Cysteine switch (Potential).
FT	CARBOHYD	547	547	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	720	720	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	764	764	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	782	782	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	945	945	N-linked (GlcNAc. . .) (Potential).

FT	DISULFID	379	462	By similarity.
FT	DISULFID	417	446	By similarity.
FT	DISULFID	571	608	By similarity.
FT	DISULFID	575	613	By similarity.
FT	DISULFID	586	598	By similarity.
FT	CONFLICT	21	21	I -> V (in Ref. 2).
FT	CONFLICT	26	31	KFRSSQ -> RSRGSL (in Ref. 2).
FT	CONFLICT	49	49	V -> A (in Ref. 2).
FT	CONFLICT	72	72	R -> P (in Ref. 2).
FT	CONFLICT	79	79	L -> TR (in Ref. 2).
FT	CONFLICT	249	249	R -> G (in Ref. 2).
FT	CONFLICT	262	265	TMLV -> NLLK (in Ref. 2).
FT	CONFLICT	607	607	S -> F (in Ref. 2).
FT	CONFLICT	936	936	L -> V (in Ref. 2).
FT	CONFLICT	962	962	I -> T (in Ref. 2).
SQ	SEQUENCE	967 AA;	105706 MW;	F93C864F6DCDB4CF CRC64;

Query Match 47.9%; Score 2471; DB 1; Length 967;  
 Best Local Similarity 48.4%; Pred. No. 2.4e-170;  
 Matches 477; Conservative 161; Mismatches 256; Indels 92; Gaps 22;

Qy	1	MLLLGILTLAFAGRTAGG--FEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQIT	58
Db	37	LLLLASITMLLCVRGAHGRPTEEDEELVL-----PSLERARGH-----DSTTLRLD	83
Qy	59	AFQEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNAE	110
Db	84	AFGQQHLHLKLQPDGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHCFYSGTVNGD	138
Qy	111	PDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS----APAAQRNSQGA----HLL--QRR	160
Db	139	PSSAAALSLCEGVRGAFYLLQGEFFIQPAPAVATERLVPAEPKEESIAPPRFHILRRRRR	198
Qy	161	GVPGGPSGD-----PTSRCGVASGWNPAILRALDPYKPRRAGFGESRRRRSGRAKRFV	214
Db	199	GSGGAKCGVMDEETLPTSNSGRESQNTPDQWPLRNP--TPQGAG---KPTGPGSIRKKRFV	254
Qy	215	SIPRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNPINIVVKVLLLRD	274
Db	255	SSPRYVETMLVADQSMADFHGSGLKHYLLTLFSVAARFYKHPSIRNSISLVVKILVIYE	314
Qy	275	RDSGPKVTGNAALTLRNFCAWQKKNKVSDDKHEPYWDTAILFTRQDLCGATTCDTLGMAD	334
Db	315	EQKGPEVTSNAALTLRNFCWQKQHNSPSDRDEHYDTAILFTRQDLCGSHTCDTLGMAD	374
Qy	335	VGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFGKLRANHMMSPTL	394
Db	375	VGTVCDPKRSVIEDDGLQAFTTAHELGHVFNMPHDDAKHCASFNGVSGDSHLMASML	434
Qy	395	IQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTSLQQCELAFGV	454
Db	435	SSLDHSQPWSPCSAYMVTSLFDNGHGECMDKPQNPILPSDLPGLTYDANRQCQFTFGE	494
Qy	455	GSKPCP-YMQYCTKLWCTGKAKGQMVQCTRHFPWADGTSCGEGKCLKGACVERHNLNKH	513
Db	495	ESTHCPDAASTCSTLWCTGTSGGLLVCQTKHFPWADGTSCGEGKWCVSGKCVNKTDM-KH	553
Qy	514	---RVDGSAKWDPYGPSCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLEPC	570
Db	554	FATPVHGSWGPWGPWGDSCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNIEDC	613
Qy	571	PSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYV	630
Db	614	PDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGYFFV	672
Qy	631	LAPKVVDGTLCSPDSTSVQVQKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFT	690
Db	673	LQPKVVDGTPCSPDSTSVQVQKQKAGCDRIIDSKKKFDKCGVCGGNGSTCKKISGTVT	732
Qy	691	KPMHGFNFVVAIPAGASSIDIRQGRYKGLIGDDNYLALKNSQKYLNLNGHFVVSVERDL	750



DR GO; GO:0006508; P:proteolysis; IEA.  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR010294; ADAM\_spacer1.  
 DR InterPro; IPR013273; ADAM\_TS.  
 DR InterPro; IPR013274; ADAM\_TS1.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR Pfam; PF05986; ADAM\_spacer1; 1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP\_1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Integrin.  
 SQ SEQUENCE 967 AA; 105648 MW; C6349B5D8CBFEA24 CRC64;

Query Match 47.8%; Score 2466; DB 2; Length 967;  
 Best Local Similarity 48.3%; Pred. No. 5.6e-170;  
 Matches 477; Conservative 160; Mismatches 255; Indels 96; Gaps 22;

Qy	1	MLLLGILTLAFAGRTAGG--FEPEREVVPIRLDPDINGRRYYWRGPEDSGDQG--LIFQ	56
		:    : :       : :    :   :	
Db	37	LLLLASITMLLCVRGAHGRPTEDEELVLP-----SLERARGHGSTTLR	81
Qy	57	ITAFQEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVN	108
		:    : :            : : :   :	
Db	82	LDAFGQQLHLKLQPD SGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHC FYSGTVN	136
Qy	109	AEPD SFAAVSLCGGLRGAFGYRGA EYVISPLPNAS----APAAQRNSQGA----HLL--Q	158
		:                 :     :     :   :   :	
Db	137	GDPSSAAALSLCEGVRGAFY LQGEFFI QPAPAVATERLVPAPKPKEESIAPPRFHILRRR	196
Qy	159	RRGVPGGPGSD-----PTSRCGVASGWNPAILRALDPYKPRRAGFGESRRRRSGRAKR	212
		:   :   :	
Db	197	RRSGGAKCGVMDEETLPTSN SGRSQNTPDQWPLRNP-TPQGAG---KPTGPGSIRKKR	252
Qy	213	FVSI PRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNPINIVVVKVLLL	272
		:     :     :     :     :     :     :     :	
Db	253	FVSSPRYVETMLVADQSMADFHGSGLKH YLLTLFSVAARFYKHPSIRNSISLVVVKILVI	312
Qy	273	RDRDSGPKVTGNAALTLRNFCAWQKKNKVSDKHPEYWDTAILFTRQDLCGATTCDTLGM	332
		:   :                 :   :   :                 :	
Db	313	YEEQKGPEVTSNAALTLRNFC SWQKQHNSPSDRDPEHYDTAILFTRQDLCGSHTCDTLGM	372
Qy	333	ADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFGKLRANHMMSP	392
		:                 :                 :     :   :	
Db	373	ADVGTVCDP SRSCSVIEDDGLQA AFTTAHELGHVFNMPHDDAKHCASFNGVSGDSHLMAS	432
Qy	393	TLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTLSQQCELA	452
		: :       :       :   :             :   :   :	
Db	433	MLSSL DHSQPWSPCSAYMVT SFLDNHGECLMDKPQNPIKLPSDLPGTLYDANRQCQFTF	492
Qy	453	GVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKLC LKGACVERHNLN	511
		:       :   :                 :     : :	
Db	493	GEESTHCPDAASTCSTLWCTGTS GGLLVCQTKHFPWADGTSCGEGKWC VSGKCVNKTDM-	551
Qy	512	KH---RVDGSWAKWD PYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLE	568
		:                 :                 :       :	
Db	552	KHFATPVHGSWGPWGPWGD SRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNIE	611
Qy	569	PCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDCKK LICRANGTG YF	628
		: :   :           : : :       :     :	
Db	612	DCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGYF	670



OM protein - protein search, using sw model

Run on: August 5, 2006, 00:04:34 ; Search time 198 Seconds  
(without alignments)  
2193.715 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*  
10: geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5164	100.0	950	4 AAG62299	Aag62299 Human met
2	5156	99.8	950	5 AAE22541	Aae22541 Human pro
3	5156	99.8	950	6 ABU09520	Abu09520 Human pro
4	5156	99.8	950	8 ADQ88214	Adq88214 Human 655
5	5149	99.7	950	6 ABR40092	Abr40092 Human ADA
6	5111	99.0	952	5 AAU74751	Aau74751 Human pro
7	4915	95.2	924	5 ABP70062	Abp70062 Human NOV
8	4856.5	94.0	928	5 AAU72899	Aau72899 Human met
9	4242.5	82.2	823	6 ABU08383	Abu08383 Human mat
10	3957.5	76.6	755	5 ABP70063	Abp70063 Human NOV
11	2497.5	48.4	505	3 AAB21257	Aab21257 Rat-metal
12	2482.5	48.1	968	6 ABU08387	Abu08387 Murine ma
13	2480.5	48.0	949	7 ADG72483	Adg72483 Human agg
14	2480.5	48.0	949	9 ADZ21014	Adz21014 Human agg
15	2480.5	48.0	950	2 AAY49501	Aay49501 Human MET
16	2480.5	48.0	950	4 AAB73549	Aab73549 Human ADA
17	2480.5	48.0	950	4 AAB50002	Aab50002 Human MET
18	2480.5	48.0	950	8 ADO20218	Ado20218 Human PRO
19	2480.5	48.0	950	8 ADQ39942	Adq39942 Human myo
20	2480.5	48.0	950	9 ADZ21015	Adz21015 Human agg
21	2480.5	48.0	950	9 AEC01581	Aec01581 Human ADA

22	2480.5	48.0	950	9	AED07146	Aed07146	Respirato
23	2480.5	48.0	967	2	AAW80285	Aaw80285	Human int
24	2480.5	48.0	967	8	ADQ39940	Adq39940	Human myo
25	2480.5	48.0	967	8	ADQ39941	Adq39941	Human myo
26	2480.5	48.0	967	9	ADY54944	Ady54944	Chronic v
27	2480.5	48.0	968	4	AAB50011	Aab50011	Protein;
28	2480.5	48.0	999	9	AED74423	Aed74423	Human pla
29	2479.5	48.0	967	2	AAY04142	Aay04142	Human Tan
30	2476.5	48.0	967	8	ADR14133	Adr14133	Human NF-
31	2458.5	47.6	967	2	AAW78189	Aaw78189	Human sec
32	2458.5	47.6	967	6	ADA57139	Ada57139	Human sec
33	2458.5	47.6	967	6	ADA41003	Ada41003	Human sec
34	2458.5	47.6	967	7	ADB91631	Adb91631	Human sec
35	2458.5	47.6	967	7	ADC74267	Adc74267	Human sec
36	2458.5	47.6	967	7	ADD37948	Add37948	Human sec
37	2445.5	47.4	950	3	AAY53899	Aay53899	Amino aci
38	2322	45.0	896	3	AAB21265	Aab21265	Mouse met
39	2274	44.0	727	2	AAW78435	Aaw78435	Human ADA
40	2136.5	41.4	890	8	ADK70513	Adk70513	Respirato
41	2135.5	41.4	890	2	AAY49502	Aay49502	Human MET
42	2135.5	41.4	890	4	AAB50003	Aab50003	Human MET
43	2134.5	41.3	924	8	ADX68264	Adx68264	Plant ful
44	2130.5	41.3	890	6	ABP96306	Abp96306	Human ADA
45	2130.5	41.3	890	9	AED90232	Aed90232	Human ADA

#### ALIGNMENTS

##### RESULT 1

AAG62299

ID AAG62299 standard; protein; 950 AA.

XX

AC AAG62299;

XX

DT 23-AUG-2001 (first entry)

XX

DE Human metalloprotease MDTS6 protein.

XX

KW Metalloprotease; human; aggrecanase; joint disease; osteoarthritis;  
KW osteopathic; antiarthritic.

XX

OS Homo sapiens.

XX

PN WO200134785-A1.

XX

PD 17-MAY-2001.

XX

PF 10-NOV-2000; 2000WO-JP007917.

XX

PR 11-NOV-1999; 99JP-00321740.

PR 16-MAY-2000; 2000JP-00144020.

XX

PA (YAMA ) YAMANOUCHI PHARM CO LTD.

PA (KAZU-) KAZUSA DNA RES INST.

XX

PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;

XX

DR WPI; 2001-343602/36.

DR N-PSDB; AAH41003.

XX

PT Metalloprotease with aggrecanase activity for treating joint diseases  
PT especially osteoarthritis.

XX

PS Claim 1; Page 56-60; 85pp; Japanese.

XX

CC This invention relates to a metalloprotease with aggrecanase activity.

CC The invention includes protein and DNA sequences of the metalloprotease,

CC vectors containing the DNA, host cells transformed by the vectors, and

CC antibodies directed against the metalloprotease. The antibodies, protein

CC and DNA sequences can be used in the treatment and prevention of joint

CC diseases, particularly osteoarthritis. The treatment may result in  
CC osteopathic and antiarthritic activity. The present sequence represents  
CC the metalloprotease of the invention termed MDTs6  
XX  
SQ Sequence 950 AA;

Query Match 100.0%; Score 5164; DB 4; Length 950;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MLLLGILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      |||
Db      1 MLLLGILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Qy     61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
      |||
Db     61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180
      |||
Db    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180

Qy    181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFSVIPRYVETLVVADESMVKFHGADLEH 240
      |||
Db    181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFSVIPRYVETLVVADESMVKFHGADLEH 240

Qy    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTNRNFCAWQKKLN 300
      |||
Db    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTNRNFCAWQKKLN 300

Qy    301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360
      |||
Db    301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360

Qy    361 HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420
      |||
Db    361 HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420

Qy    421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV 480
      |||
Db    421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV 480

Qy    481 QTRHPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDYPGPCSRTCGGGVQLAR 540
      |||
Db    481 QTRHPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDYPGPCSRTCGGGVQLAR 540

Qy    541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
      |||
Db    541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600

Qy    601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVQKCIKAGCD 660
      |||
Db    601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVQKCIKAGCD 660

Qy    661 GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
      |||
Db    661 GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720

Qy    721 GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPILPLT 780
      |||
Db    721 GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPILPLT 780

Qy    781 VEVLVSGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP 840
      |||
Db    781 VEVLVSGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP 840

Qy    841 ARWVAGSWGPCSASCGSLQKRAVDCRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900
      |||
Db    841 ARWVAGSWGPCSASCGSLQKRAVDCRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900

Qy    901 AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 950
```



|||||  
Db 901 AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950

RESULT 2

AAE22541

ID AAE22541 standard; protein; 950 AA.

XX

AC AAE22541;

XX

DT 26-JUL-2002 (first entry)

XX

DE Human protease #2.

XX

KW Human; novel human protein; NHP; protease; biological disorder; obesity;  
KW high blood pressure; arthritis; connective tissue disorder; infertility;  
KW gene therapy; enzyme.

XX

OS Homo sapiens.

XX

PN WO200226949-A2.

XX

PD 04-APR-2002.

XX

PF 27-SEP-2001; 2001WO-US030350.

XX

PR 29-SEP-2000; 2000US-0236689P.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Friddle CJ, Hilbun E;

XX

DR WPI; 2002-372123/40.

DR N-PSDB; AAD35569.

XX

PT Novel nucleic acid encoding a human protease, useful as a hybridization  
PT probe for screening libraries and assessing gene expression patterns.

XX

PS Claim 6; Page 36-38; 41pp; English.

XX

CC The present sequence is novel human protein (NHP), human protease. NHPs  
CC share structural similarity with animal proteases particularly zinc  
CC metalloproteases. Sequences of the invention are useful in therapeutic,  
CC diagnostic and pharmacogenomic applications. NHP polynucleotides are used  
CC as hybridisation probes for screening libraries and assessing gene  
CC expression patterns. They can also be used for treating related  
CC biological disorders such as obesity, high blood pressure, arthritis,  
CC connective tissue disorders and infertility. They are also used in gene  
CC therapy

XX

SQ Sequence 950 AA;

Query Match 99.8%; Score 5156; DB 5; Length 950;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
|||  
Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
  
Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
|||  
Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
  
Qy 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
|||  
Db 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
  
Qy 181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240  
|||  
Db 181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTC GGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTC GGGVQLAR	540
Qy	541	RQCTNPPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	660
Db	601	VAWVPKYSGVSPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Qy	721	GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTA VESLQASRP ILEPLT	780
Db	721	GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTA VESLQASRP ILEPLT	780
Qy	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSH PKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSH PKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVDCRGSAGQRTVPACDA AHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCGSGLQKRAVDCRGSAGQRTVPACDA AHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	950

RESULT 3  
ABU09520

ID ABU09520 standard; protein; 950 AA.

XX

AC ABU09520;

XX

DT 30-JUN-2003 (first entry)

XX

DE Human protease of the metalloprotease family.

XX

KW Human; SNP; chromosome 11; protease; metalloprotease; cancer;  
KW disintegrin; metalloprotease with thrombospondin motifs-1; cachexia;  
KW cytostatic; immunomodulator; single nucleotide polymorphism.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 220

FT /note= "Encoded by GCG in the cDNA appearing as  
FT ABX95684"

FT Misc-difference 566

```

FT          /note= "May be Lys as the result of a single nucleotide
FT          polymorphism"
FT Misc-difference 623
FT          /note= "Encoded by AGT (cDNA) or ATG (gene)"
XX
PN US2002086400-A1.
XX
PD 04-JUL-2002.
XX
PF 21-DEC-2000; 2000US-00741151.
XX
PR 06-DEC-2000; 2000US-0251398P.
XX
PA (ZHUS/) ZHU S.
PA (GUEG/) GUEGLER K.
PA (WEBS/) WEBSTER M.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Zhu S, Guegler K, Webster M, Di Francesco V, Beasley EM;
XX
DR WPI; 2003-401333/38.
DR N-PSDB; ABX95684, ABX95685.
XX
PT Novel human protease protein related to metalloprotease subfamily and
PT nucleic acid molecule encoding the protein for diagnosing, treating
PT disease or condition mediated by the protease protein e.g. cancer,
PT cachexia.
XX
PS Claim 1; Fig 2; 64pp; English.
XX
CC The invention relates to an isolated human protease peptide (a member of
CC the disintegrin/metalloprotease with thrombospondin motifs-1 family), an
CC allelic variant or orthologue encoded by a nucleic acid molecule that
CC hybridises under stringent conditions to the opposite strand of the cDNA
CC and gene appearing as (ABX95684 and ABX95685) or a fragment of having 10
CC contiguous amino acids. Also included are the encoding nucleic acids
CC (including their complements, allelic variants, orthologues or
CC fragments), an isolated antibody that selectively binds to the protease,
CC a gene chip comprising the nucleic acids, a transgenic non-human animal
CC comprising the nucleic acids, vectors and host cells producing the
CC protein. The protease is useful for identifying a modulator of the
CC expression of the protease or binds to it. The protease and the nucleic
CC acids are useful for treating, preventing and/or diagnosing disorders
CC such as cancer and cachexia. The protease and the nucleic acids are
CC further useful as a query sequence to perform a search against sequence
CC databases to identify other family members or related sequences. The
CC nucleic acids are useful as primers and probes for e.g. in situ
CC hybridisation or chromosomal localisation, and for the synthesis of
CC antisense molecules. The antibody useful to isolate, purify and detect
CC the presence of the protease in cells tissues. The gene encoding the
CC protease is located on human chromosome 11. The present sequence
CC represents the human protease
XX
SQ Sequence 950 AA;

Query Match          99.8%; Score 5156; DB 6; Length 950;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
        |||
Db      1 MLLLGILTLAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Qy      61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
        |||
Db      61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy      121 GGLRGAFGYRGA EYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180
        |||
Db      121 GGLRGAFGYRGA EYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180

```

Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI	PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI	PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT	LRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT	LRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMC	DKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMC	DKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420	
Db	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420	
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMC	480	
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMC	480	
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540	
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540	
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600	
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600	
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCD	660	
Db	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQKCIKAGCD	660	
Qy	661	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720	
Db	661	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720	
Qy	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVERSLQASRP	780	
Db	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVERSLQASRP	780	
Qy	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSHPKDPGRGPSVLHNSVLSLSNQVEQPDDRPP	840	
Db	781	VEVLSVGKMTPPRVRSFYLPKEPREDKSSHPKDPGRGPSVLHNSVLSLSNQVEQPDDRPP	840	
Qy	841	ARWVAGSWGPCASCSGLQKRAVDRCGSAGQRTVPACDAHRPVETQACGEPCPTWELS	900	
Db	841	ARWVAGSWGPCASCSGLQKRAVDRCGSAGQRTVPACDAHRPVETQACGEPCPTWELS	900	
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950	
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950	

# RESULT 4

ADQ88214

ID ADQ88214 standard; protein; 950 AA.

XX

AC ADQ88214;

XX

DT 21-OCT-2004 (first entry)

XX

DE Human 65552 protein, a metalloprotease disintegrin 15 protein ADAMTS15.

XX

KW human; cardiovascular disorder; thrombotic disorder;

KW differential expression; gene therapy; aberrant vascularisation;

KW atherosclerosis; thrombosis; coronary artery disease; hyperlipidaemia;

KW dyslipidaemia; high blood pressure; heart failure; cardiant;

KW thrombolytic; anticoagulant; antilipaemic; hypotensive; cardiant;

KW metalloprotease disintegrin 15; ADAMTS15.

XX

OS Homo sapiens.  
 XX  
 PN WO2004063340-A2.  
 XX  
 PD 29-JUL-2004.  
 XX  
 PF 13-JAN-2004; 2004WO-US000393.  
 XX  
 PR 13-JAN-2003; 2003US-0439683P.  
 PR 05-FEB-2003; 2003US-0445216P.  
 PR 18-FEB-2003; 2003US-0448036P.  
 PR 12-MAR-2003; 2003US-0454189P.  
 PR 25-MAR-2003; 2003US-0457541P.  
 PR 29-APR-2003; 2003US-0466411P.  
 PR 08-MAY-2003; 2003US-0469041P.  
 PR 10-JUN-2003; 2003US-0477414P.  
 PR 13-JUN-2003; 2003US-0478560P.  
 PR 24-JUL-2003; 2003US-0489772P.  
 PR 28-JUL-2003; 2003US-0490660P.  
 PR 03-SEP-2003; 2003US-0499838P.  
 PR 22-SEP-2003; 2003US-0504786P.  
 PR 24-SEP-2003; 2003US-0505570P.  
 PR 17-OCT-2003; 2003US-0512418P.  
 PR 27-OCT-2003; 2003US-0514660P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Stagliano NE, Healy A, Acton SL, Galvin KM, Donoghue MA;  
 PI Rogrigue-Way A, Tomlinson JE;  
 XX  
 DR WPI; 2004-553729/53.  
 DR N-PSDB; ADQ88213.  
 XX  
 PT Identifying a compound for treating a cardiovascular or thrombotic  
 PT disorder by combining a compound to be tested with e.g., a 9380, 9462,  
 PT 8701 or 2419 polypeptide or with a host cell expressing the polypeptide  
 PT and detecting the binding.  
 XX  
 PS Claim 1; SEQ ID NO 56; 512pp; English.  
 XX  
 CC This invention relates to a novel compound that is capable of treating a  
 CC cardiovascular or thrombotic disorder. Specifically, it refers to the  
 CC identification of nucleic acid molecules, and the encoded proteins  
 CC thereof, which are differentially expressed in cardiovascular disease  
 CC states relative to their normal expression in non-diseased tissue. The  
 CC present invention describes test compounds (i.e. small molecules,  
 CC peptides or antibodies) that can bind to and modulate the activity of  
 CC these differentially expressed membrane-bound polypeptides, where binding  
 CC is detected by a competition binding assay, immunoassay or yeast two-  
 CC hybrid assay. Accordingly, pharmaceutical compositions can be developed  
 CC and used via gene therapy to treat aberrant vascularisation,  
 CC atherosclerosis, thrombosis, coronary artery disease, hyperlipidaemia,  
 CC dyslipidaemia, high blood pressure or heart failure. As such, they  
 CC exhibit cardiant, thrombolytic, anticoagulant, antilipaemic, hypotensive  
 CC and cardiant activities. This polypeptide sequence is a human protein  
 CC that is differentially expressed in a patient with a cardiovascular  
 CC disorder, given in an exemplification of the invention.  
 XX  
 SQ Sequence 950 AA;

Query Match 99.8%; Score 5156; DB 8; Length 950;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
 |||||  
 Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
 Qy 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
 |||||  
 Db 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIIPRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIIPRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Qy	481	QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQKCIKAGCD	660
Db	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720
Qy	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVERSLQASRPILEPLT	780
Db	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVERSLQASRPILEPLT	780
Qy	781	VEVLSVGKMTPPRVRYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVRYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCGSLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCGSLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950

RESULT 5

ABR40092

ID ABR40092 standard; protein; 950 AA.

XX

AC ABR40092;

XX

DT 30-JUN-2003 (first entry)

XX

DE Human ADAMTS-15.

XX

KW Human; cytostatic; antiarthritic; analgesic; antiinflammatory; leukaemia;

KW immunosuppressive; ADAMTS; disintegrin; metalloprotease; thrombospondin;

KW ADAMTS-15; ADAMTS-16; ADAMTS-17; ADAMTS-18; ADAMTS-19; wound healing;

KW extracellular matrix; cancer; metastasis; arthritis; pain; coagulation;  
KW immune disorder; inflammation; immune response; angiogenesis; rheumatism;  
KW embryo implantation; foetal development; nerve degeneration.  
XX  
OS Homo sapiens.  
XX  
PN WO2003027282-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-SEP-2002; 2002WO-JP009771.  
XX  
PR 24-SEP-2001; 2001ES-00002165.  
PR 24-SEP-2001; 2001ES-00002166.  
PR 24-SEP-2001; 2001ES-00002167.  
PR 25-SEP-2001; 2001ES-00002192.  
PR 25-SEP-2001; 2001ES-00002193.  
XX  
PA (DAII-) DAIICHI FINE CHEM CO LTD.  
PA (UYOV-) UNIV OVIEDO.  
XX  
PI Cal Miguel S, Obaya Gonzalez AJ, Llamazares Prada M;  
PI Garabaya Fernandez C, Lopez-Otin C;  
XX  
DR WPI; 2003-313640/30.  
DR N-PSDB; ACC00342.  
XX  
PT ADAMTS family zinc metalloproteases for treatment and investigation of  
PT metalloprotease related disorders including cancer and arthritis.  
XX  
PS Claim 3; Page 113-118; 169pp; Japanese.  
XX  
CC The present invention relates to novel human ADAMTS (A disintegrin and  
CC metalloprotease domain-thrombospondin domain) family proteins: ADAMTS-15,  
CC ADAMTS-16, ADAMTS-17, ADAMTS-18 and ADAMTS-19 (ACC00342-ACC00346 and  
CC ABR40092-ABR40096). The ADAMTS proteins and their coding sequences are  
CC useful for the treatment, prevention, diagnosis and investigation of  
CC disorders associated with the extracellular matrix, such as cancer,  
CC cancer metastasis, leukaemia, arthritis, pain, immune disorders and  
CC inflammation; investigation of related processes such as immune response,  
CC angiogenesis, coagulation, wound healing, embryo implantation, foetal  
CC development, cancer proliferation and infiltration, metastasis,  
CC rheumatism and nerve degeneration  
XX  
SQ Sequence 950 AA;

Query Match 99.7%; Score 5149; DB 6; Length 950;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 948; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1 MLLLGILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      |||
Db      1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Qy     61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
      |||
Db     61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180
      |||
Db    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180

Qy    181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240
      |||
Db    181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300
      |||
Db    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

Qy    301 KVS DKHPEYWDTAILFTRQDL CGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360

```

Db	301		KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361		HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361		HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421		DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	480
Db	421		DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	480
Qy	481		QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481		QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541		RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541		RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601		VAWVPKYSVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD	660
Db	601		VAWVPKYSVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD	660
Qy	661		GNLGSKKRFDKCGVCGGDNKSCKKVITGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661		GNLGSKKRFDKCGVCGGDNKSCKKVITGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Qy	721		GDDNYLALKNSQKYLNGHFVVSVERDLVVGSLRLRYSGTGAVESLQASRPILPLT	780
Db	721		GDDNYLALKNSQKYLNGHFVVSVERDLVVGSLRLRYSGTGAVESLQASRPILPLT	780
Qy	781		VEVLSVGKMTPPRVYSFYLPKEPREDKSSHFKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781		VEVLSVGKMTPPRVYSFYLPKEPREDKSSHFKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841		ARWVAGSWGPCSASCSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	841		ARWVAGSWGPCSASCSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Qy	901		AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950
Db	901		AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950

RESULT 6

AAU74751

ID AAU74751 standard; protein; 952 AA.

XX

AC AAU74751;

XX

DT 09-APR-2002 (first entry)

XX

DE Human protease PRTS-11 protein sequence.

XX

KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
 KW cell proliferative disorder; developmental disorder; epilepsy;  
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
 KW reproductive disorder; endometriosis.

XX

OS Homo sapiens.

XX

PN WO200198468-A2.

XX

PD 27-DEC-2001.

XX

PF 13-JUN-2001; 2001WO-US019178.

XX

PR 16-JUN-2000; 2000US-0212336P.

PR 22-JUN-2000; 2000US-0213955P.



PR 29-JUN-2000; 2000US-0215396P.  
PR 07-JUL-2000; 2000US-0216821P.  
PR 14-JUL-2000; 2000US-0218946P.

PA (INCY-) INCYTE GENOMICS INC.

PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA:

DR WPT: 2002-090437/12.

PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.

PS Claim 1: Page 144-146: 177pp: English.

XX  
SO Sequence 952 AA:

Query Match 99.0%; Score 5111; DB 5; Length 952;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 946; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Dh 1 MLLLGILT LAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Dh 61 OEDFYLHLTPDAOFLAPAFSTEHLGVPLOGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Db 121 GGLRGAFGYRGA EYVISPLPNASAPAAQRNSOGAHL LQRRGVPGGPGDPTSRCGVASGW 180

Db 181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFSVSI PRYVETLVVADESMVKFHGADLEH 240

Db 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

Db 301 KVSDKHPEYWDTAILFTRQDLGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360

Db 361 HELGHVFNMPHDNVKVCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420

Db	421	DCLLDQPSKPISLPEDLPGASYTLSSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMC	480
Qy	481	QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDYPGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDYPGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQGKCIKAGC	659
Db	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGC	660
Qy	660	DGNLGSKKRFDKCGVCGGDNKSKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGL	719
Db	661	DGNLGSKKRFDKCGVCGGDNKSKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGL	720
Qy	720	IGDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPILPL	779
Db	721	IGDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPILPL	780
Qy	780	TVEVLVSGKMTPPRVYSFYLPKEPREDKSSHPKDPR-GPSVLHNSVLSSLNQVEQPDDR	838
Db	781	TVEVLVSGKMTPPRVYSFYLPKEPREDKSSHPHPRGGPSVLHNSVLSSLNQVEQPDDR	840
Qy	839	PPARWVAGSWGPCASCGSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCTWE	898
Db	841	PPARWVAGSWGPCASCGSGLQKRAVDWRGSAGQRTVPACDAAHRPVETQACGEPCTWE	900
Qy	899	LSAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950
Db	901	LSAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	952

RESULT 7

ABP70062

ID ABP70062 standard; protein; 924 AA.

XX

AC ABP70062;

XX

DT 27-JAN-2003 (first entry)

XX

DE Human NOV5a.

XX

KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;  
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;  
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
 KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;  
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
 KW metabolic syndrome X; wasting disorder; cell differentiation;  
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis.

XX

OS Homo sapiens.

XX

PN WO200272771-A2.

XX

PD 19-SEP-2002.

XX

PF 08-MAR-2002; 2002WO-US007288.

XX

PR 08-MAR-2001; 2001US-0274101P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 09-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 08-MAR-2002; 2002US-00093463.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;  
PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;  
PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;  
PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;  
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;  
PI Zhong M;

XX

DR WPI; 2002-732824/79.

DR N-PSDB; ABV99340.

XX

PT New NOVX polypeptides and polynucleotides, useful for preventing,

PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
PT disorders, and asthma.

XX

PS Claim 1; Page 103; 619pp; English.

XX

CC The present invention relates to new isolated proteins (NOVX) and their  
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is  
CC any number from 1 to 48. The NOVX proteins and coding sequences are  
CC useful in the manufacture of a medicament for treating a syndrome  
CC associated with a human disease, preferably a NOVX-associated disorder.  
CC The NOVX coding sequences and proteins are useful for treating,  
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
CC disease, immune disorders, haematopoietic disorders, cardiovascular  
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
CC disturbances associated with obesity, metabolic syndrome X or wasting  
CC disorders associated with chronic diseases or various cancers. The NOVX  
CC coding sequences and proteins may also be used as targets for the  
CC identification of small molecules that modulate or inhibit e.g.  
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
CC wound healing and angiogenesis, in gene therapy, in generation of  
CC antibodies that bind immunospecifically to NOVX substances for use in  
CC therapeutic or diagnostic methods

XX

SQ Sequence 924 AA;

Query Match 95.2%; Score 4915; DB 5; Length 924;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 913; Conservative 2; Mismatches 6; Indels 32; Gaps 3;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
Qy 61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
Db 61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
Qy 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW 180  
Db 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW 180  
Qy 181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240  
Db 181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240  
Qy 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300  
Db 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300  
Qy 301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360  
Db 301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360  
Qy 361 HELGHVFNMPHDNVKVCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420  
Db 361 HELGHVFNMPHDNVKVCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420  
Qy 421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC 480  
Db 421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC 480  
Qy 481 QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQ 537  
Db 481 QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHSSSQVDGSWAKWDPYGPCSRTCGGGVQ 540  
Qy 538 LARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRL 597  
Db 541 LARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPS--SGKSFREEQCEAFNGYNHSTNRL 598

Qy 598 TLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVQVQKCIKA 657  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 599 TLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVQVQKCIKA 658

Qy 658 GCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYK 717  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| : ||||||||||||||||||  
 Db 659 GCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLLSPARHGYNFVVAIPAGASSIDIRQRYK 718

Qy 718 GLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLRLRYSGTGTAVESLQASRPIL 777  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 719 GLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLRLRYSGTGTAVESLQASRPIL 778

Qy 778 PLTVEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDPRGSPVLHNSVLSLSNQVEQPDD 837  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 779 PLTVEVLSVGKMTPPRVYSFYLPKEPREDKSSH----- 812

Qy 838 RPPARWVAGSWGPCSASCSGLQKRAVDCRGSAGQRTVPACDAHRPVETQACGEPCPTW 897  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 813 -PPARWVAGSWGPCSASCSGLQKRAVDWRGSAGQRTVPACDAHRPVETQACGEPCPTW 871

Qy 898 ELSAWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 872 ELSAWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 924

RESULT 8

AAU72899

ID AAU72899 standard; protein; 928 AA.

XX

AC AAU72899;

XX

DT 26-FEB-2002 (first entry)

XX

DE Human metalloprotease partial protein sequence #11.

XX

KW Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;  
 KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;  
 KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;  
 KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;  
 KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;  
 KW lung; prostrate; cervical; brain; ovarian; bladder; kidney; pain;  
 KW immune-related disease; cardiovascular disease; neuronal disease;  
 KW migraine; sexual dysfunction; mood disorder; attention disorder;  
 KW cognition disorder; hypotension; hypertension; psychotic disorder;  
 KW dyskinesia; metabolic disorder; inflammatory disorder.

XX

OS Homo sapiens.

XX

PN WO200183782-A2.

XX

PD 08-NOV-2001.

XX

PF 04-MAY-2001; 2001WO-US014431.

XX

PR 04-MAY-2000; 2000US-0201879P.

XX

PA (SUGE-) SUGEN INC.

XX

PI Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
 PI Payne V;

XX

DR WPI; 2002-041502/05.

DR

N-PSDB; AAS97182.

XX

PT Novel protease polypeptide useful for screening for substances that may  
 PT be used to treat, e.g., cancers, immune-related diseases, cardiovascular  
 PT disease, migraine, pain, psychotic and inflammatory disorders.

XX

PS Claim 28; Fig 2G; 232pp; English.

XX

CC The invention relates to an isolated, enriched, or purified protease  
 CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to  
 CC screen for substances (S) that may modulate its activity. Administering S  
 CC (which modulates protease activity in vitro) may be used to treat a  
 CC disease or disorder selected from cancers (e.g., of tissues, of blood or  
 CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,  
 CC brain, ovarian, bladder or kidney), immune-related diseases and  
 CC disorders, cardiovascular disease, brain or neuronal-associated diseases  
 CC (e.g., central or peripheral nervous system diseases, migraine, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, hypotension, hypertension, psychotic disorders, neurological  
 CC disorders and dyskinesias), metabolic disorders and inflammatory  
 CC disorders. (I) may also be useful as a diagnostic tool for a disease or  
 CC disorder such as those above. AAU72876-AAU72910 represent human protease  
 CC amino acid sequences of the invention  
 XX  
 SQ Sequence 928 AA;

Query Match 94.0%; Score 4856.5; DB 5; Length 928;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 901; Conservative 0; Mismatches 1; Indels 49; Gaps 2;

Qy	1	MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	26	MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	85
Qy	61	QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	86	QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	145
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	180
Db	146	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	205
Qy	181	NPAILRALDPYKPRRACFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	206	NPAILRALDPYKPRRACFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	265
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	266	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	325
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	326	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	385
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	386	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	445
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSSQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	446	DCLLDQPSKPISLPEDLPGASYTLSSQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	505
Qy	481	QTRHPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	506	QTRHPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	565
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	566	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	625
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQGKCIKAGC	659
Db	626	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGC	685
Qy	660	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGL	719
Db	686	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGL	745
Qy	720	IGDDNYLALKNSQGGYLLNGHFVSAVERDLVVKGSLLRYSCTGTAVESLQASRPILPL	779

```

Db      746  |||||  IGDDNYLALKNSQGYLLNGHFVVSAVERDLVVGSLRLRYSGTGTAVESLQASRP----- 800
Qy      780  TVEVLVSVGKMTPPRVRYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRP 839
Db      801  -----NSVLSLSNQVEQPDDRP 817
Qy      840  |||||  PARWVAGSWGPCSASCGSGLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCTWEL 899
Db      818  |||||  PARWVAGSWGPCSASCGSGLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCTWEL 877
Qy      900  |||||  SAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 950
Db      878  |||||  SAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 928

```

RESULT 9

ABU08383

ID ABU08383 standard; protein; 823 AA.

XX

AC ABU08383;

XX

DT 05-JUN-2003 (first entry)

XX

DE Human matrix metalloproteinase (MMP) 65552.

XX

KW Human; matrix metalloproteinase; MMP 65552; carcinoma; sarcoma;  
 KW cellular proliferation disorder; differentiation disorder; leukaemia;  
 KW lung cancer; breast cancer; hormonal disorder; diabetes mellitus;  
 KW thyroid disorder; fertility disorder; hepatic disorder; schizophrenia;  
 KW Wilson's disease; Gaucher's disease; neurological disorder; ischaemia;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; asthma;  
 KW inflammatory disorder; Crohn's disease; immune disorder; hypertension;  
 KW autoimmune disease; multiple sclerosis; psoriasis; allergy; arthritis;  
 KW cardiovascular disorder; atherosclerosis; heart failure; bone disorder;  
 KW motility disorder; developmental disorder; skeletal disorder; obesity;  
 KW osteoporosis; osteomalacia; pain disorder; metabolic disorder; cachexia;  
 KW anorexia nervosa; blood disorder; clotting disorder; thrombocytopaenia;  
 KW cytostatic; antidiabetic; antithyroid; antiinfertility; hepatotropic;  
 KW nootropic; neuroprotective; neuroleptic; anticonvulsant; vasotropic;  
 KW antiinflammatory; immunosuppressive; antipsoriatic; antiasthmatic;  
 KW antiallergic; antiarthritic; antiarteriosclerotic; hypotensive; anaemia;  
 KW cardiant; osteopathic; analgesic; anorectic; antidepressant; haemostatic;  
 KW immunomodulator; enzyme.

XX

OS Homo sapiens.

XX

PN US2002197703-A1.

XX

PD 26-DEC-2002.

XX

PF 05-JUN-2002; 2002US-00163316.

XX

PR 13-JUN-2001; 2001US-0297863P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Kapeller-Libermann R;

XX

DR WPI; 2003-352751/33.

DR

N-PSDB; ABX94139.

XX

PT New human matrix metalloproteinase nucleic acid and polypeptide  
 PT molecules, designated 65552, useful for diagnosing, preventing or  
 PT treating cancers, diabetes, multiple sclerosis, asthma, obesity,  
 PT hypertension, pain or psoriasis.

XX

PS Claim 17; Page 42-44; 55pp; English.

XX

CC The present invention relates to the isolation of a novel human matrix  
 CC metalloproteinase (MMP), and the polynucleotide sequence encoding it. The

CC MMP 65552 polynucleotide and polypeptide sequences are useful for  
 CC diagnosing, preventing, alleviating or treating conditions associated  
 CC with aberrant expression or activity of 65552 nucleic acids or  
 CC polypeptides, e.g. cellular proliferation and/or differentiation  
 CC disorders (e.g. carcinoma, sarcoma, leukaemia, lung or breast cancer),  
 CC hormonal disorders (e.g. diabetes mellitus, thyroid disorders, fertility  
 CC disorders), hepatic disorders (e.g. Wilson's disease, Gaucher's disease),  
 CC neurological disorders (e.g. schizophrenia, ischaemia, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease), inflammatory  
 CC disorders (e.g. Crohn's disease), immune disorders (e.g. autoimmune  
 CC disease, multiple sclerosis, psoriasis, asthma, allergy, arthritis),  
 CC cardiovascular disorders (e.g. atherosclerosis, hypertension, heart  
 CC failure), motility disorders, developmental disorders, skeletal or bone  
 CC disorders (e.g. osteoporosis, osteomalacia), pain disorders, metabolic  
 CC disorders (e.g. obesity, anorexia nervosa, cachexia), and blood or  
 CC clotting disorders (e.g. anaemia or thrombocytopaenia). The present  
 CC sequence represents human MMP 65552

XX

SQ Sequence 823 AA;

Query Match 82.2%; Score 4242.5; DB 6; Length 823;  
 Best Local Similarity 96.9%; Pred. No. 0;  
 Matches 791; Conservative 0; Mismatches 2; Indels 23; Gaps 1;

Qy	1	MLLLGILTLAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGILTLAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGGVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRSRRSSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRSRRSSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC	480
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHRR-----VDG	517
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHRRPTDIISPQQLLLRLPNGLHTTQVDG	540
Qy	518	SWAKWDPYPGCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGK	577
Db	541	SWAKWDPYPGCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGK	600
Qy	578	SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD	637
Db	601	SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD	660
Qy	638	GTLCSBDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYN	697
Db	661	GTLCSBDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYN	720



Qy 698 FVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSL 757  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 FVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSL 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 758 RYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPR 793  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 RYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPR 816

RESULT 10

ABP70063

ID ABP70063 standard; protein; 755 AA.

XX

AC ABP70063;

XX

DT 27-JAN-2003 (first entry)

XX

DE Human NOV5b.

XX

KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;  
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;  
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;  
 KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;  
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;  
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;  
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;  
 KW metabolic syndrome X; wasting disorder; cell differentiation;  
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis.

XX

OS Homo sapiens.

XX

PN WO200272771-A2.

XX

PD 19-SEP-2002.

XX

PF 08-MAR-2002; 2002WO-US007288.

XX

PR 08-MAR-2001; 2001US-0274101P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 09-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275579P.

PR 13-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.

PR 16-MAR-2001; 2001US-0276776P.

PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.

PR 20-MAR-2001; 2001US-0277321P.

PR 20-MAR-2001; 2001US-0277327P.

PR 20-MAR-2001; 2001US-0277338P.

PR 21-MAR-2001; 2001US-0277791P.

PR 22-MAR-2001; 2001US-0277833P.

PR 23-MAR-2001; 2001US-0278152P.

PR 26-MAR-2001; 2001US-0278894P.

PR 27-MAR-2001; 2001US-0278999P.

PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.

PR 30-MAR-2001; 2001US-0279995P.

PR 30-MAR-2001; 2001US-0280233P.

PR 02-APR-2001; 2001US-0280802P.

PR 02-APR-2001; 2001US-0280822P.

PR 02-APR-2001; 2001US-0280900P.

PR 04-APR-2001; 2001US-0281194P.

PR 13-APR-2001; 2001US-0283675P.

PR 30-APR-2001; 2001US-0287424P.

PR 02-MAY-2001; 2001US-0288066P.

PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 08-MAR-2002; 2002US-00093463.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;  
PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;  
PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;  
PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;  
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;  
PI Zhong M;

XX

DR WPI; 2002-732824/79.

DR N-PSDB; ABV99341.

XX

PT New NOVX polypeptides and polynucleotides, useful for preventing,  
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
PT disorders, and asthma.

XX

PS Claim 1; Page 104; 619pp; English.

XX

CC The present invention relates to new isolated proteins (NOVX) and their  
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is  
CC any number from 1 to 48. The NOVX proteins and coding sequences are  
CC useful in the manufacture of a medicament for treating a syndrome  
CC associated with a human disease, preferably a NOVX-associated disorder.  
CC The NOVX coding sequences and proteins are useful for treating,  
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
CC disease, immune disorders, haematopoietic disorders, cardiovascular  
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
CC disturbances associated with obesity, metabolic syndrome X or wasting  
CC disorders associated with chronic diseases or various cancers. The NOVX  
CC coding sequences and proteins may also be used as targets for the  
CC identification of small molecules that modulate or inhibit e.g.  
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,  
CC wound healing and angiogenesis, in gene therapy, in generation of  
CC antibodies that bind immunospecifically to NOVX substances for use in  
CC therapeutic or diagnostic methods

XX

SQ Sequence 755 AA;

Query Match 76.6%; Score 3957.5; DB 5; Length 755;  
Best Local Similarity 79.4%; Pred. No. 1.5e-298;  
Matches 754; Conservative 0; Mismatches 1; Indels 195; Gaps 1;

```
Qy      1 MLLLGILTAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      |||
Db      1 MLLLGILTAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Qy     61 QEDFYLHLLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
      |||
Db     61 QEDFYLHLLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180
      |||
Db    121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180

Qy    181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240
      |||
Db    181 NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN 300
      |||
Db    241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN 300

Qy    301 KVSDKHPEYWDTAILFTRQDLCGATTCTDGLMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360
      |||
Db    301 KVSDKHPEYWDTAILFTRQ----- 319

Qy    361 HELGHVFNMPHDNVKVEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420
      |||
Db    320 ----- 319

Qy    421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGMVC 480
      |||
Db    320 ----- 319

Qy    481 QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR 540
      |||
Db    320 -----VDGSWAKWDPYGPCSRTCGGGVQLAR 345

Qy    541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
      |||
Db    346 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 405

Qy    601 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD 660
      |||
Db    406 VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD 465

Qy    661 GNLGSKKRFDKCGVCGGDNKSCKKV TGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI 720
      |||
Db    466 GNLGSKKRFDKCGVCGGDNKSCKKV TGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI 525

Qy    721 GDDNYLALKNSQGKYLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT 780
      |||
Db    526 GDDNYLALKNSQGKYLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT 585

Qy    781 VEVL SVGKMT PPRV RYSFYLPKEPREDKSSH PKDPRG P SVLHNSVLSLSNQVEQPDDRPP 840
      |||
Db    586 VEVL SVGKMT PPRV RYSFYLPKEPREDKSSH PKDPRG P SVLHNSVLSLSNQVEQPDDRPP 645

Qy    841 ARWVAGSWGPCSASCGSLQKRAVD CRGSAGQRTVPACDA AHRPVETQACGEPCPTWELS 900
      |||
Db    646 ARWVAGSWGPCSASCGSLQKRAVD CRGSAGQRTVPACDA AHRPVETQACGEPCPTWELS 705

Qy    901 AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC NLHRKPQELDFCVLRPC 950
      |||
Db    706 AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC NLHRKPQELDFCVLRPC 755
```

RESULT 11

AAB21257

ID AAB21257 standard; protein; 505 AA.

XX

AC AAB21257;

XX

DT 23-FEB-2001 (first entry)

XX

DE Rat metalloproteinase ADAMTS-5.

XX

KW Rat; ADAMTS-5; metalloproteinase; ADAM;

KW a disintegrin and metalloproteinase domain; thrombospondin domain;

KW vaccine; nootropic; neuroprotective; antiparkinsonian; cerebroprotective;

KW cytostatic; antiarthritic; immunosuppressive; Alzheimer's disease;

KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;

KW brain tumour; brain injury.

XX

OS Rattus norvegicus.

XX

PN W0200053774-A2.

XX

PD 14-SEP-2000.

XX

PR 08-MAR-2000; 2000WO-US006237.

XX

PR 08-MAR-1999; 99US-00264585.

XX

PA (NEUR-) NEUROCRINE BIOSCIENCES INC.

XX

PI Kelner GS, Clark M, Maki RA;

XX

DR WPI; 2000-594326/56.

DR

N-PSDB; AAA95827.

XX

PT Polynucleotide encoding novel members of a disintegrin, metalloproteinase

PT and thrombospondin domain protein family used to prevent and treat

PT Alzheimer's disease, cancer and autoimmune diseases.

XX

PS Claim 12; Fig 14; 129pp; English.

XX

CC The present sequence is rat metalloproteinase ADAMTS-5. The ADAMTS family

CC of proteins is closely related to the ADAM (A Disintegrin and

CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a

CC thrombospondin domain in addition to the disintegrin and

CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are

CC useful for the manufacture of medicaments for treating conditions

CC associated with neuroinflammation and/or neurodegeneration, such as

CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful

CC for treating conditions associated with cell proliferation, cell

CC migration, inflammation and/or angiogenesis, such as cancer, arthritis

CC and autoimmune diseases. They can be used to treat patients afflicted

CC with an invasive tumour, a brain tumour or brain injury

XX

SQ Sequence 505 AA;

Query Match 48.4%; Score 2497.5; DB 3; Length 505;

Best Local Similarity 93.5%; Pred. No. 3.4e-185;

Matches 462; Conservative 8; Mismatches 21; Indels 3; Gaps 2;

Qy 89 QGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLCGGLRGAFGYGA EYVISPLPNASAPAAQ 148

Db 13 QRLTGSSLDLRLRCFYSGYVNAEPDSFAAVSLCGGLRGAFGYGA EYVISPLPNTSAPEAQ 72

Qy 149 RNSQGAHLLQRRGVPVGGPSGDPTSRCGVASGWNPAILRALDPYKPRRAGFGESRSRRRSRG 208

Db 73 RNSQGAHLLQRRGAPVGGPSGDPTSRCGVASGWNPAILRALDPYKPRRTGVGESHNRRRRSRG 132

Qy 209 RAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVVVK 268

Db 133 RAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPINIVVVK 192

QY	269	VLLLRDRDSGPKVTGNAALTLRNFCAWQKLNKVS	DKHPEYWDTAILFTRQDL	CGATTCD	328
Db	193	VLLLGDRDTGPKVTGNAALTLRNFCAWQKLNKVS	DKHPEYWDTAILFTRQDL	CGATTCD	252
QY	329	TLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHEL	GHVFNMPHDNVKVCEEV	FGKLRANH	388
Db	253	TLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHEL	GHVFNMPHDNVKVCEEV	FGKLRANH	312
QY	389	MMSPTLIQIDRANPWSACSAAIITDFLDSGHGDC	LLDQPSKPISLPEDLP	GASYT	448
Db	313	MMSPTLIQIDRANPWSACSAAIITDFLDSGHGDC	LLDQPSKPITLPEDLP	PGTSY	372
QY	449	ELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	CQTRHFPWADGTSCGEG	KLCLKGAC	508
Db	373	ELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	CQTRHFPWADGTSCGEG	KFCLKGAC	432
QY	509	NLNKHRVDGWSAKWDPYGPCSRTCGGGVQLARR	--QCTNPTPANGGKY	CEGVRVKYR	566
Db	433	NPNKYRVDGPWAKWEPYGPCSRTCGGGAQLARR	QVQATLPLP-TGGKY	CEGVRVKYR	491
QY	567	LEPCPSSASGKSFR			580
Db	492	LEPCPSSASGKSFR			505

# RESULT 12

ABU08387

ID ABU08387 standard; protein; 968 AA.

XX

AC ABU08387;

XX

DT 05-JUN-2003 (first entry)

XX

DE Murine matrix metalloproteinase.

XX

KW Murine; matrix metalloproteinase; MMP; carcinoma; sarcoma; mouse;  
KW cellular proliferation disorder; differentiation disorder; leukaemia;  
KW lung cancer; breast cancer; hormonal disorder; diabetes mellitus;  
KW thyroid disorder; fertility disorder; hepatic disorder; schizophrenia;  
KW Wilson's disease; Gaucher's disease; neurological disorder; ischaemia;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; asthma;  
KW inflammatory disorder; Crohn's disease; immune disorder; hypertension;  
KW autoimmune disease; multiple sclerosis; psoriasis; allergy; arthritis;  
KW cardiovascular disorder; atherosclerosis; heart failure; bone disorder;  
KW motility disorder; developmental disorder; skeletal disorder; obesity;  
KW osteoporosis; osteomalacia; pain disorder; metabolic disorder; cachexia;  
KW anorexia nervosa; blood disorder; clotting disorder; thrombocytopaenia;  
KW cytostatic; antidiabetic; antithyroid; antiinfertility; hepatotropic;  
KW nootropic; neuroprotective; neuroleptic; anticonvulsant; vasotropic;  
KW antiinflammatory; immunosuppressive; antipsoriatic; antiasthmatic;  
KW antiallergic; antiarthritic; antiarteriosclerotic; hypotensive; anaemia;  
KW cardiant; osteopathic; analgesic; anorectic; antidepressant; haemostatic;  
KW immunomodulator; enzyme.

XX

OS Mus musculus.

XX

PN US2002197703-A1.

XX

PD 26-DEC-2002.

XX

PF 05-JUN-2002; 2002US-00163316.

XX

PR 13-JUN-2001; 2001US-0297863P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Kapeller-Libermann R;

XX

DR WPI; 2003-352751/33.

XX

PT New human matrix metalloproteinase nucleic acid and polypeptide

PT molecules, designated 65552, useful for diagnosing, preventing or  
PT treating cancers, diabetes, multiple sclerosis, asthma, obesity,  
PT hypertension, pain or psoriasis.

XX

PS Disclosure; Page 48-50; 55pp; English.

XX

CC The present invention relates to the isolation of a novel human matrix  
CC metalloproteinase (MMP), and the polynucleotide sequence encoding it. The  
CC MMP 65552 polynucleotide and polypeptide sequences are useful for  
CC diagnosing, preventing, alleviating or treating conditions associated  
CC with aberrant expression or activity of 65552 nucleic acids or  
CC polypeptides, e.g. cellular proliferation and/or differentiation  
CC disorders (e.g. carcinoma, sarcoma, leukaemia, lung or breast cancer),  
CC hormonal disorders (e.g. diabetes mellitus, thyroid disorders, fertility  
CC disorders), hepatic disorders (e.g. Wilson's disease, Gaucher's disease),  
CC neurological disorders (e.g. schizophrenia, ischaemia, Alzheimer's  
CC disease, Parkinson's disease, Huntington's disease), inflammatory  
CC disorders (e.g. Crohn's disease), immune disorders (e.g. autoimmune  
CC disease, multiple sclerosis, psoriasis, asthma, allergy, arthritis),  
CC cardiovascular disorders (e.g. atherosclerosis, hypertension, heart  
CC failure), motility disorders, developmental disorders, skeletal or bone  
CC disorders (e.g. osteoporosis, osteomalacia), pain disorders, metabolic  
CC disorders (e.g. obesity, anorexia nervosa, cachexia), and blood or  
CC clotting disorders (e.g. anaemia or thrombocytopaenia). The present  
CC sequence represents a murine MMP

XX

SQ Sequence 968 AA;

Query Match 48.1%; Score 2482.5; DB 6; Length 968;  
Best Local Similarity 48.0%; Pred. No. 1.2e-183;  
Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

Qy 1 MLLLGILTLAGRTAGG--FEPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57  
Db 37 LLLLASITMLLCARGAHGRPTEEDEELVLP-SLE-----RAPGHDSTTTRL--RL 83

Qy 58 TAFQEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGG-----SDLRRCFYSGDVNA 109  
Db 84 DAFGQQLHLKLQPDGFLAPGFTLQTV----GRSPGSEAQHLDPDGLAHCFYSGTVNG 138

Qy 110 EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS----APAAQRNSQGA----HLLQRRG 161  
Db 139 DPGSAAALSLCEGVRGAFYLGQEEFFIQPAPGVATERLAPAVPEEESARPQFHILRRR~ 197

Qy 162 VPGGPSGDPTSRGCVASGWNPAILRALDPYKPRRAGFGESRSRRR----- 206  
Db 198 ---RRGSGGAKCGVMD-----DETLPSTSDSRPESQNRNQWVVRDPTPDAGKP 243

Qy 207 ---SGRAKRFVSIPIRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI 262  
Db 244 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHLYLLTFSVAARFYKHPISIRNSI 303

Qy 263 NIVVVVKVLLLRDRDSGPKVTGNAALTLRNFCWQKKLNKVSDKHPEYWDTAILFTRQDLC 322  
Db 304 SLVVVKILVIYEEQKGPEVTSNAALTLRNFCNWQKQHNSPSDRDPEHYDTAILFTRQDLC 363

Qy 323 GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCCEVFG 382  
Db 364 GSHTCDTLGMADVGTVCDPSPRSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKHCASLNG 423

Qy 383 KLRANHMSPTLIQIDRANPWSACSAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442  
Db 424 VTGDSHLMASMLSSLDHSQPWSPCSAYMVTSLDNHGECLMDKPQNPIKLPSDLPGLTY 483

Qy 443 TLSQQCELAFGVSGSKPCP-YMQYCTKLWCTGKAKGQMVQTRHPWADGTSCGEGKCLK 501  
Db 484 DANRQCQFTFGEEKHCPDAASTCTTLWCTGTSGGLLVCTKHPWADGTSCGEGKWCVS 543

Qy 502 GACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPPTANGGKYCEGV 558  
Db 544 GKCYNKTDN-KHFATPVHGSWGPWGPWGDSCSRTCGGGVQYTMRECDNPVKNKGKYCEGK 602

```

Qy      559 RVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSQVSPRDKCKL 618
Db      603 RVRYSRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 661

Qy      619 ICRANGTGYFYVLAPKVVDGTLTCLSPDSTSVQVQGGKCIKAGCDGNLGSKKRFDKCGVCGGD 678
Db      662 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGGKAGCDRIIDSKKKFDKCGVCGGN 721

Qy      679 NKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGYLLN 738
Db      722 GSTCKKMSGIIVTSTRPGYHDIVTIPAGATNIEVKHRNQGRSRNNGSFLAIRAADGTIYN 781

Qy      739 GHFVVSAYERDLVVKGSLRLRYSGTGTAVESLQASRPILPLTVEVLSVGKMTPPRVRYSF 798
Db      782 GNFTLTSTLEQDLTYKGTVLRYSGSSAALERIRSFSPLEPLTIQVLMVGHALRPKIKFTY 841

Qy      799 YLPKEPREDKSSHPKDPGPSVLHNSVLNQLVEQPDPRPPARWVAGSWGPCASCSGSG 858
Db      842 FMKKKTES-----FNAIPTFS-----EWVIEEWGECSTCGSG 874

Qy      859 LQKRAVDCRGSAQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915
Db      875 WQRRVVCQRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCSKTCGKGYKK 931

Qy      916 RSLKCVGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950
Db      932 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 967

```

# RESULT 13

ADG72483

ID ADG72483 standard; protein; 949 AA.

XX

AC ADG72483;

XX

DT 11-MAR-2004 (first entry)

XX

DE Human aggrecanase (ADAMTS-1).

XX

KW Human; aggrecanase; ADAMTS-1; rheumatoid arthritis; osteoarthritis;

KW antiarthritic; osteopathic; enzyme.

XX

OS Homo sapiens.

XX

PN US6649377-B1.

XX

PD 18-NOV-2003.

XX

PF 09-MAY-2000; 2000US-00568559.

XX

PR 10-MAY-1999; 99US-0133343P.

XX

PA (SYNT ) SYNTEX USA LLC.

XX

PI Allard JD, Ileller RA, Klonowski P, Vanwart HE;

XX

DR WPI; 2003-896162/82.

DR

N-PSDB; ADG72482.

XX

PT New nucleic acid present in other than its natural environment, useful  
PT for preparing a composition for diagnosing or treating diseases  
PT associated with aggrecanase activity, e.g. rheumatoid arthritis or  
PT osteoarthritis.

XX

PS Disclosure; SEQ ID NO 2; 26pp; English.

XX

CC The present invention relates to the isolation of human aggrecanase  
CC (ADAMTS-1), and the polynucleotide sequence encoding it. Also disclosed  
CC are is a method of producing aggrecanase in vitro. The polynucleotide  
CC sequence encoding aggrecanase is useful for preparing a composition for

CC diagnosing or treating diseases associated with aggrecanase activity e.g.  
CC rheumatoid arthritis or osteoarthritis. The present sequence represents  
CC human aggrecanase.

Sequence 949 AA;

Query Match 48.0%; Score 2480.5; DB 7; Length 949;  
Best Local Similarity 48.4%; Pred. No. 1.7e-183;  
Matches 482; Conservative 153; Mismatches 250; Indels 111; Gaps 22;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQLIFQITAF 60  
: ||| || : | : ||| | : | | | : ||  
Db 18 LLLLLAAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTTRLRLHAF 63

Qy           61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF   114  
               :   | | | : | | | : :: |                   ||                   : ||   | | | | | : |  
Db           64 DOOLDLELRPDSSFLAPGFTLONVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA   118

Qy            115 AAVSLCGGLRGAFGYRGA EYVISPLPNAS---APAAQRNSQGA----HLLQRRGVPGGPS 167  
             ||:||| |:|||| | | |||| | | || | |||||  
Db            119 AALSLCEGVRGAFYLLGEAYFIOPLPAA SERLATAAPEKPPAPLQFHLLRR-----NRQ 173

Qy 168 GDPTSRCGV-----ASGWNPAILRALDPYKPRRAGFGESR 202  
|| ||| : | | | | :  
Db 174 GDVGTCGVVDDEPRPTGKAETEDEGTEGEDEGAQWS-----PQDPALQGVGQP- 224

Qy 203 SRRRSGRAKRFVSIPIRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI 262  
: | | |||| | ||||::|||:| :||: |:||||| : ||||:|||| | :  
Db 225 TGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTFSVAARLYKHPSIRNSV 284

Qy 263 NIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVSDDKHPEYWDTAILFTQDLC 322  
::|||:::| ||::| ||||| |||: | ||: |:| ||||| |||  
Db 285 SLVVVKILVIHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTQDLC 344

Qy 323 GATTCDTLGMADVGTMCDPKRSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVC EEVFG 382  
| : ||||| : || ||||| : ||||| : || : |  
Db 345 GSQTCDTLGMADVGTVCDPKRSVIEDDGLQAFTTAHELGHVFNMPHDDAKQCASLNG 404

Qy 383 KLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442  
: : ||| : | : | : ||| ||| : || |||| : ||| : || : ||  
Db 405 VNQDSHMMASMLSNLDHSQPWSPCSAYMITSFGLNGHGECMDKPQNPIQLPGDLPGTSY 464

Qy 443 TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAQGMCVQTRHFPWADGTSCGEGKCLKL 501  
::||: || || | : ||| : : ||||| ||||| ||| :  
Db 465 DANRQCQFTFGEDSKHCDDAASTCSTLWCTGTSGGVLCVQTKEFPWADGTSCGEGKCIN 524

Qy 502 GACVERHNLNKH---RVDGSWAKWDYPGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV 558  
| | | : : | | | | : | | | | | | | | : | | | | | | | |  
Db 525 GKC VNKTH-RKHFDTPFHGSGWGMGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGK 583

Qy 559 RVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAVPKYSGVSPRDKCKL 618  
||:||||| || : :||:||||| | :: : || :||:||||:|:|

Db 584 RVRYRSCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPDKRCKL 642

Qy 619 ICRANGTGYFYVLAPKVVDTGLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGD 678  
||:| | ||:| | ||||| | ||||| ||||| :|:| ||| : ||:| ||||| |:  
Db 643 ICQAKGIGYFFVLQPKVVDTGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKRFDKCGVCGGN 702

Qy 679 NKSCCKVVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQ GKYL LN 738  
: |||::| | ||: :: || ||::|:::| | : | : ::||:| : | |::|

Db 703 GSTCKKISGSVTSAPKPGYHDIITPTGATNIEVKQRNQRGSRNNGSFLAIKAADGTYILN 762

Qy 739 GHFVVSAVERDLVVGKSLRLRYSGTGTAVESLQASRPILPLTVEVL SVGKMTPPVRVYSF 798  
| : : | : |:|:: || :||| |:| :: |: |||::|:|| |::|::

Db 763 GDYTLSTLEQDIMYKGV LRYSGSSAALERIRSF SPLKEPLTIQVLT VGNALRPKIKYTY 822

Qy 799 YLPKEPREDKSSHPKDPRGFSVLHNSVLSLSNQVEQPDDRPPARWVAGSWGPCASCSGSG 858  
:: | : | : : | | | | | |  
Db 823 FVKKK---KES-----FNAIPTFS-----AWVIEEWGECSSKSELG 855

Qy 859 LQKRAVDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915



```

Db      856 WQRRLVECRDINGQ---PASECAKEVKPASTRPCADHPCPQWLGEWSSCSKTCGKGYKK 912
Qy      916 RSLKCVGHGGRLRLARDQCNLHRKPQE-LDFCVLRPC 950
Db      913 RSLKCLSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 948

```

# RESULT 14

ADZ21014

ID ADZ21014 standard; protein; 949 AA.

XX

AC ADZ21014;

XX

DT 14-JUL-2005 (first entry)

XX

DE Human aggrecanase (ADAMTS-1), SEQ ID NO:2 #1.

XX

KW Protein degradation; musculoskeletal disease; rheumatoid arthritis;  
KW osteoarthritis; gout; psoriatic arthritis; spondylosis; injury; fibrosis;  
KW pulmonary disease; inflammation; antiarthritic; osteopathic;  
KW antirheumatic; antigout; antiinflammatory; respiratory-gen.; ADAMTS-1;  
KW aggrecanase; enzyme.

XX

OS Homo sapiens.

XX

PN US2005100916-A1.

XX

PD 12-MAY-2005.

XX

PF 17-SEP-2003; 2003US-00667281.

XX

PR 10-MAY-1999; 99US-0133343P.

PR 09-MAY-2000; 2000US-00568559.

XX

PA (ALLA/) ALLARD J D.

PA (HELL/) HELLER R A.

PA (KLON/) KLONOWSKI P.

PA (VWAR/) VAN WART H E.

XX

PI Allard JD, Heller RA, Klonowski P, Van Wart HE;

XX

DR WPI; 2005-345393/35.

DR N-PSDB; ADZ21013.

XX

PT New isolated human aggrecanase useful for diagnosing or treating  
PT conditions associated with aggrecanase activity, e.g. rheumatoid  
PT arthritis, osteoarthritis, joint trauma or fibrosis.

XX

PS Claim 33; SEQ ID NO 2; 26pp; English.

XX

CC The invention relates to human aggrecanase (ADZ21014 and ADZ21015), also  
CC known as ADAMTS-1 (a disintegrin and metalloproteinase with  
CC thrombospondin motifs 1), and its fragments. The invention also discloses  
CC nucleic acids encoding human aggrecanase (ADZ21013), methods of treating  
CC disease conditions associated with aggrecanase activity, and methods of  
CC modulating aggrecanase activity in a cell. In arthritic cartilage, the  
CC major cartilage proteoglycan aggrecan is cleaved by proteases to produce  
CC 55 kD and 60 kD fragments. Aggrecanase is responsible for generation of  
CC the 60 kD fragment, cleaving aggrecan between Glu373 and Ala374 in a  
CC region located between globular domains G1 and G2 to yield a fragment  
CC with the N-terminal sequence shown in ADZ21016. Cleavage in this region  
CC leads to the loss of aggrecan and exposes type II collagen to  
CC collagenases, causing cartilage loss and the consequent development of  
CC joint disease. Human aggrecanase and nucleic acids encoding it are useful  
CC in the diagnosis or treatment of conditions associated with aggrecanase  
CC activity such as rheumatoid arthritis, osteoarthritis, infectious  
CC arthritis, gouty arthritis, psoriatic arthritis, spondylosis, sports  
CC injury, joint trauma, pulmonary disease and fibrosis. They may also be  
CC used for research purposes and in drug screening procedures. The present  
CC sequence represents human aggrecanase. Note: The present sequence is 1

CC residue shorter than the version fo SEQ ID NO:2 shown in Figure 1.  
CC Although the present sequence is designated SEQ ID NO:2 in the sequence  
CC listing and the disclosure, an aggrecanase amino acid sequence is  
CC referred to in Claim 1. However, the claim refers to the amino acid  
CC sequence as SEQ ID NO:1, which is the aggrecanase-encoding nucleic acid  
CC in the sequence listing and disclosure.

XX

SQ Sequence 949 AA;

Query Match 48.0%; Score 2480.5; DB 9; Length 949;  
Best Local Similarity 48.4%; Pred. No. 1.7e-183;  
Matches 482; Conservative 153; Mismatches 250; Indels 111; Gaps 22;

```
Qy      1 MLLLGILTAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      :|||  ||  :  |||  ||  :  |||  ||  :  |||  ||  :  |||  ||  :  |||
Db      18 LLLLLAAALLAVSDALGRPSEEEDELVVP-ELE-----RAP---GHGTTLRRLHAF 63

Qy      61 QEDFYHLHTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF 114
      :  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      64 DQQLDELRLPDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCFYSGTVNGDPSSA 118

Qy     115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQRRGVPGGPS 167
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     119 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLLRR-----NRQ 173

Qy     168 GDPTSRCGV-----ASGWNPAILRALDPYKPRRAGFGESR 202
      ||  |||  :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db     174 GDVGTCGVVDDEPRPTGKAETEDEGTEGEDEGAQWS-----PQDPALQGVGQP- 224

Qy     203 SRRRSGRAKRFVSIPIRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI 262
      :  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     225 TGTGSTRKKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPSIRNSV 284

Qy     263 NIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKNKVSDKHPEYWDTAILFTRQDLC 322
      :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     285 SLVVVKILVIHDEQKGEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLC 344

Qy     323 GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCCEEVFG 382
      ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     345 GSQTCDTLGMADVGTVCPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNG 404

Qy     383 KLRAHMMSPSTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442
      :  |||  ||  :  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     405 VNQDSHMMASMLSNLDHSQPWSPCSAYMITSFLDNGHGECMDKQNPQLPGDLPGTSY 464

Qy     443 TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGMVQCQTRHPWADGTSCGEGKCLK 501
      :|||  ||  ||  ||  |||  ||  |||  |||  |||  |||  |||  |||  |||
Db     465 DANRQQQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWCIN 524

Qy     502 GACVERHNLNKH---RVDGSAKWDPYGPCSRTCGGGVQLARRQCTNPPTANGGKYCEGV 558
      ||  ||  :  ||  |||  ||  |||  |||  |||  |||  |||  |||  |||
Db     525 GKCVNKT-HRKHFDTPFHGSWMGWPWDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGK 583

Qy     559 RVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAVWPKYSGVSPRDKCKL 618
      |||  |||  ||  :  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     584 RVRYRSCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 642

Qy     619 ICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGGKCIKAGCDGNLGSKKRFDKCGVCGGD 678
      |||  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     643 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGGCVKAGCDRIIDSKKKFDKCGVCGGN 702

Qy     679 NKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRRGYKGLIGDDNYLALKNSQGKYLNL 738
      :|||  ||  |||  ||  |||  ||  |||  ||  |||  ||  |||  ||  |||
Db     703 GSTCKKISGSVTSAPPGYHDIITPTGATNIEVKQRNQRGSRNNGSFLAIKAADGTIYILN 762

Qy     739 GHFVVSAVERDLVVKGSLRLRYSGTGTAVESLQASRPILPTVEVLSVGKMTPPRVRYSF 798
      |  :  :  |||  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db     763 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPKLEPLTIQVLTVGNALRPKIYTY 822

Qy     799 YLPKEPREDKSSHPKDPGRGPSVLHNSVLSLSNQVEQPDDRRPPARWVAGSWGPCSASCGSG 858
      :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
```

Db	823	FVKKK-----KES-----FNAIPTFS-----AWVIEEWGECskSCELg	855
Qy	859	LQKRAVDCRGsAGQRTVPACDAAH--RPVETQACGE-PCPTWELsAWSPCSKsCGRGFQR	915
		:   :          :   :	
Db	856	WQRRlVECRDINGQ--PASECAKEVKPASTRPCADHPCPQWQLGEWSSsSKTCGKGYKK	912
Qy	916	RSLKCVGHGGRLlARDQCNlHRKpQE-LDFCVLrPC	950
		:   :   :     :	
Db	913	RSLKCLSHDGGVlSHEsCDPLKKPKHfIDFCTMAEC	948

RESULT 15

```

ID      AAY49501 standard; protein; 950 AA.
XX
AC      AAY49501;
XX
DT      10-JAN-2000 (first entry)
XX
DE      Human METH1 protein.
XX
KW      Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW      cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW      angiogenesis inhibitor; abnormal wound healing; inflammation;
KW      rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW      diabetic retinopathy; macula degeneration; haemangioma; detection;
KW      arterial-venous malformation; immune deficiency.
XX
OS      Homo sapiens.
XX
PN      WO9937660-A1.
XX
PD      29-JUL-1999.
XX
PF      22-JAN-1999; 99WO-US001313.
XX
PR      23-JAN-1998; 98US-0072298P.
PR      28-AUG-1998; 98US-0098539P.
XX
PA      (IRUE/) IRUELA-ARISPE L.
PA      (HAST/) HASTINGS G A.
PA      (RUBE/) RUBEN S M.
XX
PI      Iruela-Arispe L, Hastings GA, Ruben SM;
XX
DR      WPI; 1999-590684/50.
DR      N-PSDB; AAZ32000.
XX
PT      New isolated metalloprotease thrombospondin polypeptides, useful for
PT      treating hyperproliferative disorders, cancers or autoimmune disorders.
XX
PS      Claim 10; Fig 1; 457pp; English.
XX
CC      AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC      metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC      respectively. METH1 and METH2 have been found to be potent inhibitors of
CC      angiogenesis both in vitro and in vivo. They can be used for treating
CC      cancer and other disorders related to angiogenesis including abnormal
CC      wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial
CC      bleeding disorders, diabetic retinopathy, some forms of macula
CC      degeneration, haemangiomas, and arterial-venous malformations. They may
CC      be useful in treating deficiencies or disorders of the immune system, by
CC      activating or inhibiting the proliferation, differentiation, or
CC      mobilisation (chemotaxis) of immune cells. The etiology of these immune
CC      deficiencies or disorders may be genetic, somatic, such as cancer or some
CC      autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
CC      infectious. They can also be used to treat inflammatory conditions, both
CC      chronic and acute conditions. The products can also be used for detection
CC      and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to AAY49511 represent
CC      sequences given in the exemplification of the present invention
XX

```

SQ Sequence 950 AA;

Query Match 48.0%; Score 2480.5; DB 2; Length 950;  
Best Local Similarity 48.9%; Pred. No. 1.7e-183;  
Matches 485; Conservative 154; Mismatches 251; Indels 101; Gaps 24;

```
QY      1 MLLLGILTLAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      :|||  || :  | : |||  | :  | |  | :  ||
Db      19 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RAP---GHGTTTRLRLHAF 64

QY      61 QEDFYHLTPDAQFLAPAFSTEHLG-----VPLQGLTGGSSDLRRCFYSGDVNAEPDSF 114
      :  | | ||: ||| | : :|  |  |  | ||  ||| | | : | |
Db      65 DQQLDLELRPDSSFLAPGFTLQNVGRKSGSETPLP-----ETDLAHCIFYSGTVNGDPSSA 119

QY     115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQ--RRGVPGG 165
      ||: ||| | : |||  | | | ||| ||  | |  |  |||: | : | |
Db     120 AALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAPLQFHLLRRNRQGDVGG 179

QY     166 PSG-----DPTSRC-----GVASG--WNPAILRALDPYKPRRAGFGESRSRRRS 207
      |  |  |  | :  |  |  | :  |  |  | :  |  |
Db     180 TCGVVDDEPRPTGKAETEDEDEGTEGEDEGPQWS-----PQDPALQGVGQP-TGTGS 230

QY     208 GRAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLATAARLYRHPSILNPINIVV 267
      | |||||  ||||: |||: | : ||| | : |||||  | : ||||: ||| | : |||
Db     231 IRKKRFVSSHRYVETMLVADQSMAEFHGSGLKHYLLTLFSAARLYKHPSIRNSVSLVV 290

QY     268 KVLRLRDRDSGPKVTGNAALTLRNFCWQKKLNKVS DKHPEYWDTAILFTRQDLCGATTC 327
      |: |: |  | : ||| ||||| |||: |  | : | : ||||| |||||: | |
Db     291 KILVIHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQTC 350

QY     328 DTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCCEEVFGKLRAN 387
      ||||| |||||: ||| ||||| ||||| : ||||| |||||: | | : | : :
Db     351 DTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNQDS 410

QY     388 HMMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASYTLSQQ 447
      |||: |  | : ||| | : |||: |||: |||: | | | ||| | | : |
Db     411 HMMASMLSNLDHSQPWSPCSAYMITSFLDNHGHECLMDKPQNPIQLPGDLPGTSYDANRQ 470

QY     448 CELAFGVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKLCLKGACVE 506
      | : | | | |  | : |||| | : ||||: ||||| ||||| ||| : | | |
Db     471 CQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCGEGKWCINGKCVN 530

QY     507 RHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGVRVKYR 563
      : : | |  | | | ||||| ||| : | | | ||||| ||| : | |
Db     531 KTD-RKHFDTPFHGSWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVYR 589

QY     564 SCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRAN 623
      ||||| | : |||: ||||| | : : | | : |||: |||||: |||||: |
Db     590 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAK 648

QY     624 GTGYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCDNLGSKKRFDKCGVCGGDNKSCK 683
      | : |||: | | ||||| ||||| |||||: ||||| : |||: ||||| ||| : ||
Db     649 GIGYFFVLQPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCK 708

QY     684 KVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGKYLLNGHFVV 743
      | : | | |  | : : | | |||: |||: |  | : : |||: |  | | : | :
Db     709 KISGSVTSAPKPYHDIITIPTGATNIEVKQRNQGRSRNNGSFLAIKAADGTIYILNGDYTL 768

QY     744 SAVERDLVVKGSLLRYSGTGTAVESLQASRPILPTVEVLSVGKMTPPRVRYSFYLPKE 803
      | : |||: | | : ||||| | : ||| : | : |||||: ||| : |||: |||: |
Db     769 STLEQDIMYKGVVLRYSGSSAALERIRSFSPLEPLTIQVLTGVGNALRPKIKYTYFVKKK 828

QY     804 PREDKSSHPKDPGRPSVLHNSVLSLSNQVEQPDDRPPARWAGSWGPCASCGSGLQKRA 863
      | |  |  | :  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     829 ---KES-----FNAIPTFS-----AWVIEEWGECSKSCELGWQRR 861

QY     864 VDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQRRSLKC 920
      | : | | |  | : |  | : | : | : | : | : | : | : | : | : |
Db     862 VECRDINGQ---PASECAKEVKPASTRPCADHPCPQWQLGEWSSCSKTCGKGYKKRSLKC 918

QY     921 VGHGGRLLARDQC�NLHRKPQE-LDFCVLRPC 950
```

Db           : | | : | : | : | : | : | : |  
919 LSHDGGVLSHESCDPLKKPKHFIDFCTMAEC 949

Search completed: August 5, 2006, 00:08:26  
Job time : 204 secs

OM protein - protein search, using sw model

Run on: August 5, 2006, 00:26:10 ; Search time 186 Seconds  
(without alignments)  
2365.882 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQC�LHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA\_Main:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5164	100.0	950	4	US-10-763-210-1
2	5156	99.8	950	3	US-09-741-151-2
3	5156	99.8	950	3	US-09-965-631-4
4	5156	99.8	950	4	US-10-391-364-77
5	5156	99.8	950	5	US-10-753-267-56
6	5156	99.8	950	5	US-10-961-020-4
7	5121.5	99.2	978	4	US-10-275-107-59
8	5111	99.0	952	4	US-10-311-035-11
9	4915	95.2	924	4	US-10-093-463-28
10	4242.5	82.2	823	4	US-10-163-316-2
11	3957.5	76.6	755	4	US-10-093-463-30
12	2482.5	48.1	950	3	US-09-321-987B-4
13	2482.5	48.1	951	4	US-10-381-793-3
14	2482.5	48.1	968	4	US-10-163-316-7
15	2482.5	48.1	968	4	US-10-391-364-82
16	2480.5	48.0	931	3	US-09-741-151-4
17	2480.5	48.0	949	5	US-10-667-281-2
18	2480.5	48.0	950	3	US-09-373-658-2
19	2480.5	48.0	950	3	US-09-989-687-2
20	2480.5	48.0	950	5	US-10-741-600-1605
21	2480.5	48.0	950	5	US-10-973-858-14
22	2480.5	48.0	967	4	US-10-105-929-2
23	2480.5	48.0	967	4	US-10-115-286-2
24	2480.5	48.0	967	4	US-10-757-450-2
25	2480.5	48.0	967	5	US-10-741-600-1603

26	2480.5	48.0	967	5	US-10-741-600-1604	Sequence 1604, Ap
27	2480.5	48.0	967	5	US-10-923-035-37	Sequence 37, Appl
28	2480.5	48.0	967	5	US-10-756-149-5619	Sequence 5619, Ap
29	2480.5	48.0	967	5	US-10-757-450-2	Sequence 2, Appli
30	2480.5	48.0	967	5	US-10-995-561-979	Sequence 979, App
31	2480.5	48.0	967	5	US-10-995-561-980	Sequence 980, App
32	2480.5	48.0	967	6	US-11-169-041-191	Sequence 191, App
33	2480.5	48.0	968	3	US-09-373-658-125	Sequence 125, App
34	2480.5	48.0	999	5	US-10-821-234-1251	Sequence 1251, Ap
35	2476.5	48.0	967	4	US-10-755-889-134	Sequence 134, App
36	2471.5	47.9	967	3	US-09-989-687-126	Sequence 126, App
37	2458.5	47.6	967	6	US-11-111-953-179	Sequence 179, App
38	2274	44.0	727	3	US-09-445-023A-1	Sequence 1, Appli
39	2274	44.0	727	4	US-10-097-597-1	Sequence 1, Appli
40	2274	44.0	727	4	US-10-097-580-1	Sequence 1, Appli
41	2273	44.0	727	3	US-09-445-023A-12	Sequence 12, Appl
42	2273	44.0	727	4	US-10-097-597-12	Sequence 12, Appl
43	2273	44.0	727	4	US-10-097-580-12	Sequence 12, Appl
44	2135.5	41.4	890	3	US-09-373-658-4	Sequence 4, Appli
45	2135.5	41.4	890	3	US-09-989-687-4	Sequence 4, Appli

# ALIGNMENTS

## RESULT 1

US-10-763-210-1

; Sequence 1, Application US/10763210

; Publication No. US20040142445A1

; GENERAL INFORMATION:

; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

; APPLICANT: Kazusa DNA Research Institute

; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY

; FILE REFERENCE: Q67541

; CURRENT APPLICATION NUMBER: US/10/763,210

; CURRENT FILING DATE: 2004-01-26

; PRIOR APPLICATION NUMBER: US/10/009,332

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: JPA Hei 11-321740

; PRIOR FILING DATE: 1999-11-11

; PRIOR APPLICATION NUMBER: JPA 2000-144020

; PRIOR FILING DATE: 2000-05-16

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 950

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-763-210-1

Query Match 100.0%; Score 5164; DB 4; Length 950;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300

Db	241		300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301		360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361		420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	480
Db	421		480
Qy	481	QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481		540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541		600
Qy	601	VAWVPKYSVGSPPDKCKLICRANGTYFYVLAPKVVDGTLCSPTDSTSVCVQGKCIKAGCD	660
Db	601		660
Qy	661	GNLGSKKRFDKCGVCGDNKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661		720
Qy	721	GDDNYLALKNSQGYLLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPILPLT	780
Db	721		780
Qy	781	VEVLSVGKMTPPRVRYSFYLPKEPREDKSSHFKDPRGSPVLHNSVLSLSNQVEQPDDRPP	840
Db	781		840
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVDRCGSAGQRTVPACDAHRPVETQACGEPCTWELS	900
Db	841		900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC	950
Db	901		950

RESULT 2

US-09-741-151-2

; Sequence 2, Application US/09741151

; Publication No. US20020086400A1

; GENERAL INFORMATION:

; APPLICANT: ZHU, Shiaooping et al

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001005

; CURRENT APPLICATION NUMBER: US/09/741,151

; CURRENT FILING DATE: 2000-12-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 950

; TYPE: PRT

; ORGANISM: Human

US-09-741-151-2

Query Match 99.8%; Score 5156; DB 3; Length 950;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



Qy	1	MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIIPRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIIPRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAWVPKYSGVSPRDCKKLCIRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	660
Db	601	VAWVPKYSGVSPRDCKKLCIRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGGDNKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGGDNKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Qy	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPILEPLT	780
Db	721	GDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPILEPLT	780
Qy	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCGSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQC NLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQC NLHRKPQELDFCVLRPC	950

RESULT 3

US-09-965-631-4

; Sequence 4, Application US/09965631

; Patent No. US20020115842A1

; GENERAL INFORMATION:





Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV	480
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQGCKIKAGCD	660
Db	601	VAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQGCKIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGGDNKSCKKVTLGFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGGDNKSCKKVTLGFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720
Qy	721	GDDNYLALKNSQGKYLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT	780
Db	721	GDDNYLALKNSQGKYLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT	780
Qy	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCSGSLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCSGSLQKRAVDCRGSAQRTVPACDAAHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC NLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC NLHRKPQELDFCVLRPC	950

RESULT 5

US-10-753-267-56

; Sequence 56, Application US/10753267

; Publication No. US20050037946A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Stagliano, Nancy E.

; APPLICANT: Healy, Aileen

; APPLICANT: Acton, Susan L.

; APPLICANT: Galvin, Katherine M.

; APPLICANT: Donoghue, Mary A.

; APPLICANT: Rodrigue-Way, Amelie

; APPLICANT: Tomlinson, James E.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,  
; TITLE OF INVENTION: 10653, 9235, 21668, 17794, 2210, 6169, 10102, 21061, 17662,  
; TITLE OF INVENTION: 1468, 12282, 6350, 9035, 1820, 23652, 7301, 8925, 8701,  
; TITLE OF INVENTION: 3533, 9462, 9123, 12788, 17729, 65552, 1261, 21476, 33770,  
; TITLE OF INVENTION: 9380, 2569654, 33556, 53656, 44143, 32612, 10671, 261,  
; TITLE OF INVENTION: 44570, 41922, 2552, 2417, 19319, 43969, 8921, 8993, 955,  
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,  
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419  
; FILE REFERENCE: MPI03-003PIRNOMNIM  
; CURRENT APPLICATION NUMBER: US/10/753,267  
; CURRENT FILING DATE: 2004-01-08  
; PRIOR APPLICATION NUMBER: US 60/439,683  
; PRIOR FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US 60/445,216  
; PRIOR FILING DATE: 2003-02-05  
; PRIOR APPLICATION NUMBER: US 60/448,036  
; PRIOR FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/454,189  
; PRIOR FILING DATE: 2003-03-12  
; PRIOR APPLICATION NUMBER: US 60/457,541  
; PRIOR FILING DATE: 2003-03-25  
; PRIOR APPLICATION NUMBER: US 60/466,411  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 60/469,041  
; PRIOR FILING DATE: 2003-05-08  
; PRIOR APPLICATION NUMBER: US 60/477,414  
; PRIOR FILING DATE: 2003-06-10  
; PRIOR APPLICATION NUMBER: US 60/478,560  
; PRIOR FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: US 60/489,772  
; PRIOR FILING DATE: 2003-07-24  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 56  
; LENGTH: 950  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-753-267-56

Query Match 99.8%; Score 5156; DB 5; Length 950;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
Db 1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60  
Qy 61 QEDFYHLHLPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
Db 61 QEDFYHLHLPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDSFAAVSLC 120  
Qy 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
Db 121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180  
Qy 181 NPAILRALDPYKPRRAGFGESRSRRSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH 240  
Db 181 NPAILRALDPYKPRRAGFGESRSRRSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH 240  
Qy 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCWQKKLN 300  
Db 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCWQKKLN 300  
Qy 301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360  
Db 301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360  
Qy 361 HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420

```

Db      361  |||||
Qy      421  DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV 480
Db      421  DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMV 480
Qy      481  QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR 540
Db      481  QTRHFPWADGTSCGEGKLCCKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR 540
Qy      541  RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
Db      541  RQCTNPTPANGGKYCEGVRVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
Qy      601  VAWVPKYSVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD 660
Db      601  VAWVPKYSVSPRDKCKLICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCD 660
Qy      661  GNLGSKKRFDKCGVCGDNKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Db      661  GNLGSKKRFDKCGVCGDNKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI 720
Qy      721  GDDNYLALKNSQGGKYLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPILEPLT 780
Db      721  GDDNYLALKNSQGGKYLNGHFVVSVERDLVVGSLRLRYSGTGTAVESLQASRPILEPLT 780
Qy      781  VEVLVSGKMTPPRVRYSFYLPKEPREDKSSHFKDPRGPSVLHNSVLSLSNQVEQPDDRPP 840
Db      781  VEVLVSGKMTPPRVRYSFYLPKEPREDKSSHFKDPRGPSVLHNSVLSLSNQVEQPDDRPP 840
Qy      841  ARWVAGSWGPCSASCGSGLQKRAVDGRSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900
Db      841  ARWVAGSWGPCSASCGSGLQKRAVDGRSAGQRTVPACDAAHRPVETQACGEPCPTWELS 900
Qy      901  AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950
Db      901  AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC 950

```

# RESULT 6

US-10-961-020-4

```

; Sequence 4, Application US/10961020
; Publication No. US20050090657A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0241-USA
; CURRENT APPLICATION NUMBER: US/10/961,020
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/965,631
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/236,689
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-961-020-4

```

```

Query Match          99.8%; Score 5156; DB 5; Length 950;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 949; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1  MLLLGILTLAGRTAGGFEPEREVVPRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
Db      1  MLLLGILTLAGRTAGGSEPEREVVPRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

```

Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRSRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMC	480
Db	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMC	480
Qy	481	QTRHFPWADGTSCGEGKLC LK GACVERHNLNKH RV DGSWAKWDPYGPCSRTCGGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLC LK GACVERHNLNKH RV DGSWAKWDPYGPCSRTCGGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPC PSSASGKS FREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPC PSSASGKS FREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQ GKCIKAGCD	660
Db	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQ GKCIKAGCD	660
Qy	661	GNLGSKKRFDKCGVCGGDNKSCKKV TGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Db	661	GNLGSKKRFDKCGVCGGDNKSCKKV TGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLI	720
Qy	721	GDDNYLALKNSQ GK YLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT	780
Db	721	GDDNYLALKNSQ GK YLLNGHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLT	780
Qy	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSH PKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Db	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSH PKDPRGPSVLHNSVLSLSNQVEQPDDRPP	840
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVDCRG SAGQRTVPACDA AHRPVETQACGEPCPTWELS	900
Db	841	ARWVAGSWGPCSASCGSGLQKRAVDCRG SAGQRTVPACDA AHRPVETQACGEPCPTWELS	900
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC	950
Db	901	AWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQCNLHRKPQELDFCVLRPC	950

# RESULT 7

US-10-275-107-59

; Sequence 59, Application US/10275107

; Publication No. US20040063107A1

## GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY D.

; APPLICANT: WHYTE, DAVID

; APPLICANT: SUDARSANAM, SUCHA

; APPLICANT: MANNING, GERARD

; APPLICANT: CAENEPEEL, SEAN R.

; APPLICANT: PAYNE, VILLA  
; TITLE OF INVENTION: NOVEL PROTEASES  
; FILE REFERENCE: 038602/1479  
; CURRENT APPLICATION NUMBER: US/10/275,107  
; CURRENT FILING DATE: 2003-11-03  
; PRIOR APPLICATION NUMBER: PCT/US01/14431  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/201,879  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 978  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-275-107-59

X

Query Match 99.2%; Score 5121.5; DB 4; Length 978;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 947; Conservative 0; Mismatches 3; Indels 3; Gaps 2;

Qy	1	MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	26	MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	85
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	86	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	145
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	146	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	205
Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH	240
Db	206	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSIPTYVETLVVADESMVKFHGADLEH	265
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	266	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	325
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	326	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	385
Qy	361	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	386	HELGHVFNMPHDNVKVCCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	445
Qy	421	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	446	DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	505
Qy	481	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDYPGCSRTCGGGVQLAR	540
Db	506	QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDYPGCSRTCGGGVQLAR	565
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	600
Db	566	RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA	625
Qy	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPK-VVDGTLCSPDSTSVCVQGKCIKAGC	659
Db	626	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGC	685
Qy	660	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGL	719
Db	686	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGL	745
Qy	720	IGDDNYLALKNSQGKYLNGHFVVSVERDLVVKGSLLRYSGTGTAVESLQASRPPILEPL	779



```

Db      746 IGDDNYLALKNSQGKYLLNGHFVVSVERDLVVKGSLRLRYSGTGTAVESLQASRPILPL 805
Qy      780 TVEVLVSGKMTPPRVRYSFYLPKEPREDKSSHPKDPR--GPSVLHNSVLSLSNQVEQPDD 837
      |||
Db      806 TVEVLVSGKMTPPRVRYSFYLPKEPREDKSSHPPHPRGGGPSVLHNSVLSLSNQVEQPDD 865
      |||
Qy      838 RPPARWVAGSWGPCSASCGSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTW 897
      |||
Db      866 RPPARWVAGSWGPCSASCGSGLQKRAVDRCGSAGQRTVPACDAAHRPVETQACGEPCPTW 925
      |||
Qy      898 ELSAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 950
      |||
Db      926 ELSAWSPCSKSCGRGFQRRSLKCVGHGGRLRLARDQCNLHRKPQELDFCVLRPC 978
      |||

```

# RESULT 8

US-10-311-035-11

; Sequence 11, Application US/10311035

; Publication No. US20040023243A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: YUE, Henry

; APPLICANT: ELLIOTT, Vicki

; APPLICANT: GANDHI, Ameena R.

; APPLICANT: LAL, Preeti

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: TRIBOULEY, Catherine M.

; APPLICANT: DELEGEANE, Angelo M.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: NGUYEN, Dannie B.

; APPLICANT: LEE, Ernestine A.

; APPLICANT: HAFALIA, April

; APPLICANT: KHAN, Farrah A.

; APPLICANT: CHAWLA, Narinder K.

; APPLICANT: YAO, Monique G.

; APPLICANT: LU, Dyung Aina M.

; APPLICANT: ARVIZU, Chandra S.

; APPLICANT: TANG, Y. Tom

; APPLICANT: WALSH, Roderick T.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LU, Yan

; APPLICANT: RAMKUMAR, Jayalaximi

; APPLICANT: XU, Yuming

; APPLICANT: REDDY, Roopa

; APPLICANT: DAS, Depopriya

; APPLICANT: KEARNEY, Liam

; APPLICANT: KALLICK, Deborah A.

; TITLE OF INVENTION: Proteases

; FILE REFERENCE: PI-0123 PCT

; CURRENT APPLICATION NUMBER: US/10/311,035

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946

; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PERL Program

; SEQ ID NO 11

; LENGTH: 952

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: Incyte ID No. US20040023243A1 7473089CD1

US-10-311-035-11

Query Match 99.0%; Score 5111; DB 4; Length 952;

Best Local Similarity 99.48%; Pred. No. 0;

Matches 946; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

Qy      1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      |||
Db      1 MLLLGILTLAFAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
      |||

```

Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAHEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAHEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALT LRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDL CGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDL CGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPI SLPEDLP GASYTLSQQCELA FGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	421	DCLLDQPSKPI SLPEDLP GASYTLSQQCELA FGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Qy	481	QTRHFPWADGTSCGEGKLC LKGACVERHNLNKHVRVDGSAKWDPYGPCSRTC GGGVQLAR	540
Db	481	QTRHFPWADGTSCGEGKLC LKGACVERHNLNKHVRVDGSAKWDPYGPCSRTC GGGVQLAR	540
Qy	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPC PSSASGKS FREEQCEAFNGYNHSTNRLTLA	600
Db	541	RQCTNPTPANGGKYCEGVRVKYRSCNLEPC PSSASGKS FREEQCEAFNGYNHSTNRLTLA	600
Qy	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPK-VVDGTLCS PDSTSVCVQGKCIKAGC	659
Db	601	VAVVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVVDGTLCS PDSTSVCVQGKCIKAGC	660
Qy	660	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGL	719
Db	661	DGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRYKGL	720
Qy	720	IGDDNYLALKNSQ GKYL LNHGFVVS AVERDLVVKGSLLRYSGTGTAVESLQASRP ILEPL	779
Db	721	IGDDNYLALKNSQ GKYL LNHGFVVS AVERDLVVKGSLLRYSGTGTAVESLQASRP ILEPL	780
Qy	780	TVEVL SVGKMT PPRVRY SFYLPKEPREDKSSH PKDPR-GPSVLHNSVL SLSNQVEQPDDR	838
Db	781	TVEVL SVGKMT PPRVRY SFYLPKEPREDKSSH PPHPRGGPSVLHNSVL SLSNQVEQPDDR	840
Qy	839	PPARWVAGSWGPCSASCGSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEP CPTWE	898
Db	841	PPARWVAGSWGPCSASCGSGLQKRAVD WRGSAGQRTVPACDAAHRPVETQACGEP CPTWE	900
Qy	899	LSAWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	950
Db	901	LSAWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	952

# RESULT 9

US-10-093-463-28

; Sequence 28, Application US/10093463

; Publication No. US20030208039A1

## ; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal  
; APPLICANT: Zhong, Mei  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Mezes, Peter  
; APPLICANT: Smithson, Glennnda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Vernet, Corine  
; APPLICANT: Pena, Carol  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Gorman, Linda  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Voss, Edward  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Anderson, David  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Miller, Charles  
; APPLICANT: Taupier, Raymond J. Jr.  
; TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypeptides, Nucleic  
Acids  
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.  
; FILE REFERENCE: 21402-290A (Cura 590AT)  
; CURRENT APPLICATION NUMBER: US/10/093,463  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: 60/283,675  
; PRIOR FILING DATE: 2001-04-14  
; PRIOR APPLICATION NUMBER: 60/338,092  
; PRIOR FILING DATE: 2001-12-03  
; PRIOR APPLICATION NUMBER: 60/274,281  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,101  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/325,681  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/304,354  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/279,995  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/294,899  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/287,424  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/299,027  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/309,198  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/281,194  
; PRIOR FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,849  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/330,380  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/275,235  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/288,342  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/275,578  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 370  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 28

2

Qy	1	MLLLGILTLAFAGRTAGGFEPEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Db	1	MLLLGILTLAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF	60
Qy	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Db	61	QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC	120
Qy	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Db	121	GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW	180
Qy	181	NPAILRALDPYKPRRAGFGESRRRSRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Db	181	NPAILRALDPYKPRRAGFGESRRRSRAKRFVSI PRYVETLVVADESMVKFHGADLEH	240
Qy	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Db	241	YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN	300
Qy	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Db	301	KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA	360
Qy	361	HELGHVFNMPHDNVKVC EEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Db	361	HELGHVFNMPHDNVKVC EEVFGKLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHG	420
Qy	421	DCLLDQPSKPI SLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Db	421	DCLLDQPSKPI SLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKQMVC	480
Qy	481	QTRHFPWADGTSCGEGKCLKGACVERHNLNKH---RVDGSWAKWDPYGP CSRTCGGGVQ	537
Db	481	QTRHFPWADGTSCGEGKCLKGACVERHNLNKHSSSQVDGSWAKWDPYGP CSRTCGGGVQ	540
Qy	538	LARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRL	597
Db	541	LARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSS--SGKSFREEQCEAFNGYNHSTNRL	598
Qy	598	TLAVAWVPKYSVGSVPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKA	657
Db	599	TLAVAWVPKYSVGSVPRDKCKLICRANGTYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKA	658
Qy	658	GCDGNLGSKKRFDKCGVCGGDNKSCKKV TGLFTKPMHGYNFVVAIPAGASSIDIRQGYK	717
Db	659	GCDGNLGSKKRFDKCGVCGGDNKSCKKV TGLLSPARHGYNFVVAIPAGASSIDIRQGYK	718
Qy	718	GLIGDDNYLALKNSQ GKYYLLNGHFVVS AVERDLVVKGSLLRYSGTGTAVESLQASRPILE	777
Db	719	GLIGDDNYLALKNSQ GKYYLLNGHFVVS AVERDLVVKGSLLRYSGTGTAVESLQASRPILE	778
Qy	778	PLTVEVL SVGKMT PPRVRY SFYLPKEPREDKSSHPKDP RGPSVLHNSVLSLSNQVEQPDD	837
Db	779	PLTVEVL SVGKMT PPRVRY SFYLPKEPREDKSSH-----	812
Qy	838	RPPARWVAGSWGPCSASCGSLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEP CPTW	897
Db	813	-PPARWVAGSWGPCSASCGSLQKRAVD WRGSAGQRTVPACDAAHRPVETQACGEP CPTW	871
Qy	898	ELSAWSPCSKSCGRGFQRRSLKCVGHGGRLLARDQC NLHRKPQELDFC VLRPC	950

US-10-163-316-2

: Sequence 2. Application US/10163316

: Publication No. US20020197703A1

: GENERAL INFORMATION:

; APPLICANT: Kapeller-Libermann, Rosana

: TITLE OF INVENTION: 65552. A Human Matrix Metalloproteinase and Uses

: TITLE OF INVENTION: Therefor

; FILE REFERENCE: MPI01-025P1RNM

CURRENT APPLICATION NUMBER: US/10/163,316

CURRENT FILING DATE: 2002-06-05

: PRIOR APPLICATION NUMBER: 60/297,863

: PRIOR FILING DATE: 2001-06-13

: NUMBER OF SEO ID NOS: 10

```

; SOFTWARE: FastSEO for Windows Version 4.0

```

: SEO ID NO 2

: LENGTH: 823

TYPE: PRT

: ORGANISM: Homo sapiens

US-10-163-316-2

Query Match 82.2%; Score 4242.5; DB 4; Length 823;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 791; Conservative 0; Mismatches 2; Indels 23; Gaps 1;

Qy 1 MLLLGILTLAFAGRTAGGFEPEREVVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Db 1 MLLLGILT LAFAGRTAGGSEPEREVVVPIRLDPDINGRRYYWRGPEDSGDOGLIFOITAF 60

QY 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Db 61 OEDFYLHLTPDAOFLAPAFSTEHLGVPLOGLTGGSSDLRRCFYSGGVNAEPDSFAAVSLC 120

Qy 121 GGLRGAFGYRGA EYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPSGDPTSRCGVASGW 180

Db 121 GGLRGAFGYRGA EYVISPLPNASAPAAQRNSQGAHLLQRRGVPPGGPSGDPTSRCGVASGW 180

Qy 181 NPAILRALDPYKPRRAGFGESRSRRRSRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Db 181 NPAILRALDPYKPRRAGFGESRRRSRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

Db 241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

[illegible]

Db 301 KVSDKHPEYWDTAILFTRQDLGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360

Qy 361 HELGHVFNMPHDNVKVC EEVFGKLRANHMMSP TLIQIDRANPWSACSAAIITDFLDSGHG 420

Db 361 HELGHVFNMPHDNVKVCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLD SGHG 420

Qy 421 DCLLDQPSKPISLPEDLPGASYTSLQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGMVC 480

Db 421 DCLLDQPSKPISLPEDLPGASYTLSQQCELAFGVGSKPCPYMQYCTKLWCTGKAKGQMVC 480

Qy 481 QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHRL-----VDG 517

Db 481 QTRHFPWADGTSCGEGKLCLKGACVERHNLNKHPPPTDIISPQQLLRPNGLHTTQVDG 540

Qy 518 SWAKWDPYGPCSRTC GGGVQLARRQCTNP TPANGGKYCEGVRVKYRSCNLEPCPSSASGK 577

Db 541 SWAKWDPYGPCSRTC GGGVQLARRQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGK 600

Qy 578 SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVD 637

Db 601 SFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRANGTGYFYVLAPKVVD 660

Qy 638 GTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYN 697  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 GTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYN 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 698 FVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 757  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 FVVAIPAGASSIDIRQRGYKGLIGDDNYLALKNSQGYLLNGHFVVSVERDLVVKGSLL 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 758 RYSGTGTAVESLQASRPFILEPLTVEVLSVGKMTTPR 793  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 RYSGTGTAVESLQASRPFILEPLTVEVLSVGKMTTPR 816  
 ||||||||||||||||||||||||||||||||||||||||||||

RESULT 11

US-10-093-463-30

; Sequence 30, Application US/10093463

; Publication No. US20030208039A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal

; APPLICANT: Zhong, Mei

; APPLICANT: Rastelli, Luca

; APPLICANT: Mezes, Peter

; APPLICANT: Smithson, Glennnda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gerlach, Valerie

; APPLICANT: Casman, Stacie

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gangolli, Esha

; APPLICANT: Vernet, Corine

; APPLICANT: Pena, Carol

; APPLICANT: Burgess, Catherine

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gorman, Linda

; APPLICANT: Spaderna, Steven

; APPLICANT: Voss, Edward

; APPLICANT: Malyankar, Uriel

; APPLICANT: Anderson, David

; APPLICANT: Patturajan, Meera

; APPLICANT: Miller, Charles

; APPLICANT: Taupier, Raymond J. Jr.

; TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids

; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.

; FILE REFERENCE: 21402-290A (Cura 590AT)

; CURRENT APPLICATION NUMBER: US/10/093,463

; CURRENT FILING DATE: 2002-06-24

; PRIOR APPLICATION NUMBER: 60/283,675

; PRIOR FILING DATE: 2001-04-14

; PRIOR APPLICATION NUMBER: 60/338,092

; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: 60/274,281

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: 60/274,101

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: 60/325,681

; PRIOR FILING DATE: 2001-09-27

; PRIOR APPLICATION NUMBER: 60/304,354

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/279,995

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 60/294,899

; PRIOR FILING DATE: 2001-05-31

```

; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-093-463-30

```

```

Query Match          76.6%; Score 3957.5; DB 4; Length 755;
Best Local Similarity 79.4%; Pred. No. 0;
Matches 754; Conservative 0; Mismatches 1; Indels 195; Gaps 1;

```

```

Qy      1 MLLLGILTLAGRTAGGFEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      1 MLLLGILTLAGRTAGGSEPEREVVPIRLDPDINGRRYYWRGPEDSGDQGLIFQITAF 60

Qy      61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 QEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGSSDLRRCFYSGDVNAEPDSFAAVSLC 120

Qy      121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW 180
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      121 GGLRGAFGYRGAEYVISPLPNASAPAAQRNSQGAHLLQRRGVPGGPGSDPTSRCGVASGW 180

Qy      181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      181 NPAILRALDPYKPRRAGFGESRRRRSGRAKRFVSI PRYVETLVVADESMVKFHGADLEH 240

Qy      241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      241 YLLTLLATAARLYRHPSILNPINIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLN 300

Qy      301 KVSDKHPEYWDTAILFTRQDLCGATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTA 360
        ||||||||||||||||
Db      301 KVSDKHPEYWDTAILFTRQ----- 319

Qy      361 HELGHVFNMPHDNVKCEEVFGKLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHG 420
Db      320 ----- 319

Qy      421 DCLLDQPSKPISLPEDLPGASYTLSSQCELAFGVGSKPCPYMQYCTKLWCTGKAKGMVC 480
Db      320 ----- 319

Qy      481 QTRHFPWADGTSCGEGKLCCLKGACVERHNLNKHVRDGSWAKWDPYGPCSRTCGGGVQLAR 540
        ||||||||||||||||
Db      320 -----VDGSWAKWDPYGPCSRTCGGGVQLAR 345

Qy      541 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 600
        ||||||||||||||||
Db      346 RQCTNPTPANGGKYCEGVRVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLA 405

```

Qy	601	VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	660
Db	406	VAWVPKYSVSPRDKCKLICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQGKCIKAGCD	465
Qy	661	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	720
Db	466	GNLGSKKRFDKCGVCGGDNKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQRGYKGLI	525
Qy	721	GDDNYLALKNSQGKYLNGHFVVS AVERDLVVGSLRLRYSGTGTA VESLQASRP ILEPLT	780
Db	526	GDDNYLALKNSQGKYLNGHFVVS AVERDLVVGSLRLRYSGTGTA VESLQASRP ILEPLT	585
Qy	781	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDP RGP SVLHNSVLSLSNQVEQPDDRPP	840
Db	586	VEVLSVGKMTPPRVYSFYLPKEPREDKSSHPKDP RGP SVLHNSVLSLSNQVEQPDDRPP	645
Qy	841	ARWVAGSWGPCSASCGSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	900
Db	646	ARWVAGSWGPCSASCGSGLQKRAVD CRGSAGQRTVPACDAAHRPVETQACGEPCPTWELS	705
Qy	901	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	950
Db	706	AWSPCSKSCGRGFQRRSLKCVGHGGRL LARDQCNLHRKPQELDFCVLRPC	755

RESULT 12

US-09-321-987B-4

```

; Sequence 4, Application US/09321987B
; Patent No. US20020102210A1
; GENERAL INFORMATION:
; APPLICANT: Kimble, Judith E
; APPLICANT: Blelloch, Robert H
; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
; FILE REFERENCE: 960296.95386
; CURRENT APPLICATION NUMBER: US/09/321,987B
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,170
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/129,023
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Murine
US-09-321-987B-4

```

Query Match 48.1%; Score 2482.5; DB 3; Length 950;  
 Best Local Similarity 48.0%; Pred. No. 2.6e-199;  
 Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

Qy	1	MLLLGILTLAFAGRTAGG--FEPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI	57
Db	20	LLLLASITMLLCARGAHRPTEDEELVLP-SLE-----RAPGHDSTTTRL--RL	66
Qy	58	TAFQEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA	109
Db	67	DAFGQQLHLKLQPD SGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHC FYSGTVNG	121
Qy	110	EPDSFAAVSLCGGLRGAFGYRGA EYVISPLPNAS----APAAQRNSQGA----HLLQRRG	161
Db	122	DPGSAAALSLCEGVRGAFY LQGEFFIQPAPGVATERLAPAVPEEESARPQFHILRRR-	180
Qy	162	VPGGPSGDPTSRCGVASGWNPA ILRALDPYKPRRAGFGESRSRRR-----	206
Db	181	----RRGGGAKCGVMD-----DETLP TSDSRPESQNTRNQWPVRDPTPDAGKP	226
Qy	207	----SGRAKRFVSI PRYVETLVVADES MVKFHGADLEHYLLTLLATAARLYRHPSILNPI	262



Db 227 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSLKHYLLTLFSVAARFYKHPsirnsi 286  
 Qy 263 NIVVVVLLLLRRDSGPKVTGNAALTLRNFCAWQKLNKVSdkhPEYWDTAILFTRQDLC 322  
 Db 287 SLVVVKILVIYEEQKGPEVTSNAALTLRNFCNWQKQHNSPSDRDPEHYDTAILFTRQDLC 346  
 Qy 323 GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFG 382  
 Db 347 GSHTCDTLGMADVGTVCDSRSCSVIEDDGLQAFTTAHELGHVFNMPHDDAKHCASLNG 406  
 Qy 383 KLRANHMMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442  
 Db 407 VTGDSHLMASMLSSLDHSQWSPCSAYMVSFLDNHGECLMDKPQNPIKLPSDLPGTLY 466  
 Qy 443 TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKCLK 501  
 Db 467 DANRQCQFTFGEEKHCPDAASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSCGEGKCVS 526  
 Qy 502 GACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV 558  
 Db 527 GKCvNKTDM-KHFATPVHGSWGPWGPWGDCSRTCGGGVQYTMRECDNPVKNNGGKYCEGK 585  
 Qy 559 RVKYRSCNLEPCSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRKCKL 618  
 Db 586 RVRYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 644  
 Qy 619 ICRANGTGYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGD 678  
 Db 645 TCEAKGIGYFFVLQPKVVDGTSPSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGN 704  
 Qy 679 NKSCKKVTLGFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGKYLLN 738  
 Db 705 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHRNQGRSRNNGSFLAIRAADGTIYLN 764  
 Qy 739 GHFVVSaverDLVVKGSLLRYSGTGTAVESLQASRPilePLTVEVLSVGKMTFPRVRYSF 798  
 Db 765 GNFTLSTLEQDLTYKGTVLRYSGSSAALERIRSFSPLEPLTIQVLMVGHALRPKIKFTY 824  
 Qy 799 YLPKEPREDKSSHPKDPGSPVLHNSVLSLSNQVEQPDDRPPARWAGSWGPCASCGSG 858  
 Db 825 FMKKKTES-----FNAIPTFS-----EWVIEEWGECSKTCGSG 857  
 Qy 859 LQKRAVDRCRSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915  
 Db 858 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCHWQVGDWSPCSKTCGKGYKK 914  
 Qy 916 RSLKCVGHGGRLRLARDQCNLHRKPQE-LDFCVLRPC 950  
 Db 915 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 950

RESULT 13

US-10-381-793-3

; Sequence 3, Application US/10381793

; Publication No. US20040091965A1

; GENERAL INFORMATION:

; APPLICANT: Bayer AG

; TITLE OF INVENTION: REGULATION OF HUMAN ADAM-TS-LIKE PROTEIN

; FILE REFERENCE: LIO152 Foreign Countries

; CURRENT APPLICATION NUMBER: US/10/381,793

; CURRENT FILING DATE: 2003-03-28

; PRIOR APPLICATION NUMBER: US 60/235,881

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US 60/XXX,XXX

; PRIOR FILING DATE: 2001-07-24

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 951

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-381-793-3

Query Match 48.1%; Score 2482.5; DB 4; Length 951;  
Best Local Similarity 48.0%; Pred. No. 2.6e-199;  
Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

```
Qy      1 MLLLGILTAFAGRTAGG--FEPEREVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
      :||| :|: | | | | :|:|: | : | | | :|
Db      20 LLLLASITMLLCARGAHRPTEEDEELVLP-SLE-----RAPGHDSTTTRL--RL 66

Qy      58 TAFQEDFYHLHTPDQAQFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA 109
      || : :| | ||: ||| |: : : | : || | | ||| ||
Db      67 DAFGQQLHLKLQPD SGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHC FYSGTVNG 121

Qy     110 EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS----APAAQRNSQGA----HLLQRRG 161
      :| | ||: || |: ||| :| |: | | | : || | | |:|: ||
Db     122 DPGSAAALSLCEGVRGAFYLGQEEFFIQPAPGVATERLAPAVPEEESARPQFHILRRR- 180

Qy     162 VPGGPSGDPTSRCGVASGWNPAILRALDPYKPRRAGFGESRSGR----- 206
      | :||| | | | ||:|
Db     181 ----RRGSGGAKCGVMD-----DETLPTSDSRPESQNTRNQWVPVRDPTPDAGKP 226

Qy     207 ----SGRAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI 262
      | | ||||| |||||: ||:| ||: | ||||| : || |: ||| | |
Db     227 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHYYLLTLFSAARFYKHPSIRNSI 286

Qy     263 NIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVS DKHPEYWDTAILFTRQDLC 322
      :||: ||: : ||: ||||| ||: | ||: ||: ||||| |||||
Db     287 SLVVVKILVIYEEQKGPEVTSNAALTLRNFCNWQKHNSPSDRDPEHYDTAILFTRQDLC 346

Qy     323 GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFG 382
      |: ||||| |||||: ||| ||||| ||||| ||||| ||||| :| | :|
Db     347 GSHTCDTLGMADVGTVC DPSRSCSVIEDDGLQAAF'TTAHELGHVFNMPHDDAKHCASLNG 406

Qy     383 KLRANHHMSP TLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442
      :||: | :| : ||: || ||: ||: ||: ||: || || || |||
Db     407 VTGDSHLMASMLSSL DHSQWPSPCSAYMVT SFLDNHGHECLMDKPQNPIKLPSDLPGTLY 466

Qy     443 TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGQMVQC TRHFPWADGTSCGEGKCLK 501
      :||: || || || ||||| :| :|||: ||||| ||||| ||:
Db     467 DANRQCQFTFGEESKHCPDAASTCTTLWCTGTSGGLLVQC TKHFPWADGTSCGEGKWCVS 526

Qy     502 GACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV 558
      | || : :| : ||: ||||| ||||| ||: || || ||||| |||
Db     527 GKC VNKTM-KHFATPVHGSWGPWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGK 585

Qy     559 RVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSGVSPRDCKKL 618
      ||: ||: || : :||: ||||| || : : || : ||: ||||| :|||
Db     586 RVRYSRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 644

Qy     619 ICRANGTGYFYVLAPKVVDGTL CSPDSTSVCVQGGKCIKAGCDGNLGSKKRFDKCGVCGGD 678
      | | | ||: || ||||| ||||| ||||| ||: ||||| : ||: ||||| |||
Db     645 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGGQVKAGCDRIIDSKKKFDKCGVCGGN 704

Qy     679 NKSCKKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGKYLLN 738
      :||: ||: | ||: ||||| ||: || :| : :||: ||: | |||
Db     705 GSTCKKMSGIVTSTRPGYHDI VTI PAGATNIEVKHRNQGRSRNNGSFLAIRAADGTYILN 764

Qy     739 GHFVVS AVERDLVVKGSLLRYS GTGTAVESLQASRP ILEPLTVEVLSVGKMTPPRVRYSF 798
      |: | :| : ||: || ||||| ||: ||: | ||||| || || :||: ||
Db     765 GNFTLSTLEQDLTYKGT VLRYSGSSAALERIRSF SPLKEPLTIQVLMVGHALRPKIKFTY 824

Qy     799 YLPKEPREDKSSH PKDPRGPSVLHNSVLSLSNQVEQPDDRPPARWVAGSWGPCSASCGSG 858
      :| : | : ||: || ||: || || || || |||
Db     825 FMKKKTES-----FNAIPTFS-----EWVIEEWGECSKTCGSG 857

Qy     859 LQKRAVD CRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915
      | | | || | | :| | :| : ||| :| ||||| ||: ||: ||
Db     858 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCSKTCGKGYYK 914

Qy     916 RSLKCVGHGGRL LARDQCNLHRKPQE-LDFCVLRPC 950
```

Db 915 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 950

RESULT 14

```

US-10-163-316-7
; Sequence 7, Application US/10163316
; Publication No. US20020197703A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MPI01-025PlRNM
; CURRENT APPLICATION NUMBER: US/10/163,316
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,863
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-163-316-7

```

Query Match 48.1%; Score 2482.5; DB 4; Length 968;  
Best Local Similarity 48.0%; Pred. No. 2.6e-199;  
Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

Qy	1	MLLLGILTLAFAGRTAGG--FEPEREVVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI	57
Db	37	LLLLASITMLLCARGAHRPTEEDEELVLP-SLE-----RAPGHDSTTTTL--RL	83
Qy	58	TAFQEDFYHLHTPDAQFLAPAFSTEHLGVPLQGLTGGS-----SDLRRCFYSGDVNA	109
Db	84	DAFGQQLHLKLQPDSGFLAPGFTLTQTV----GRSPGSEAQHLDPTGDLAHCIFYSGTVNG	138
Qy	110	EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQRRG	161
Db	139	DPGSAAALSLCEGVRGAFYLQGEFFIQQAPGVATERLAPAVPEEESARPQFHILRRR-	197
Qy	162	VPGGPSGDPTSRCGVASGWNPAILRALDPYKPRRAGFGESRSRRR-----	206
Db	198	----RRSGGAKCGVMD-----DETLPTSDSRPESQNTNRNQWVPVRDPTPDQAGKP	243
Qy	207	----SGRAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI	262
Db	244	SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHYLLTFSVAARFYKHPISIRNSI	303
Qy	263	NIVVVVKVLLLRDRDSDGPKVTGNAALTNRNFCAWQKKLNKVSDDKHPEYWDTAILFTRQDLC	322
Db	304	SLVVVKILVIYEEQKGPEVTSNAALTNRNFCNWQKQHNSPSDRDPEHYDTAILFTRQDLC	363
Qy	323	GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVC EEVFG	382
Db	364	GSHTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCASLNG	423
Qy	383	KLRANHHMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY	442
Db	424	VTGDSHLMASMLSSLDHSQPWSPCSAYMVSFLDNHGHCELMDKPQNPIKLP S DLP GTLY	483
Qy	443	TLSQQCELAFGVGSKPCP-YMQYCTKLWCTGKAKGQMVCQTRHFPWADGTSCGEGKLCCLK	501
Db	484	DANRQCQFTFGEESKHC PDAASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSCGEGKWCVS	543
Qy	502	GACVERHNLNKH---RVDGSAKWDPYGPSCSRTCGGGVQLARRQCTNP TPANGGKYCEGV	558
Db	544	GKCVNKTDM-KHFATPVHGSWGPWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGK	602
Qy	559	RVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSVSPRDKCKL	618

[illegible]

; ORGANISM: Mus musculus  
US-10-391-364-82

Query Match 48.1%; Score 2482.5; DB 4; Length 968;  
Best Local Similarity 48.0%; Pred. No. 2.6e-199;  
Matches 478; Conservative 158; Mismatches 249; Indels 111; Gaps 22;

```
Qy      1 MLLLGILTAFAGRTAGG--FEPEREVVPIRLDPDINGRRYYWRGP-EDSGDQGLIFQI 57
      :||| :|:  || | | : |:| | : | | | | :| :
Db      37 LLLLASITMLLCARGAHRPTEEDEELVLP-SLE-----RAPGHDSTTTRL--RL 83

Qy      58 TAFQEDFYLHLPDAQFLAPAFSTEHLGVPLQGLTGGs-----SDLRRCFYSGDVNA 109
      || : :| | | | | | | : : : | : || | | | | | |
Db      84 DAFGQQLHLKLQPD SGFLAPGFTLQTV-----GRSPGSEAQHLDPTGDLAHCFYSGTVNG 138

Qy     110 EPDSFAAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQRRG 161
      :| | | | | | | | | | | : | | | | : | | | | | | |
Db     139 DPGSAAALSLCEGVRGAFYLGQEEFFIQPAPGVATERLAPAVPEEESSARPOFHILRRR- 197

Qy     162 VPGGPGSDPTSRCGVASGWNPAILRALDPYKPRRAGFGESRSRR----- 206
      | : ||| | | | | | | | | | | | | | | | |
Db     198 ---RRGSGGAKCGVMD-----DETLPSTDSRPESQNTRNQWPVRDPTPDAGKP 243

Qy     207 ---SGRAKRFVSI PRYVETLVVADESMVKFHGADLEHYLLTLLATAARLYRHPSILNPI 262
      | | | | | | | | | | | | | | | | | | | | | |
Db     244 SGPGSIRKKRFVSSPRYVETMLVADQSMADFHGSGLKHYLLTLFVSVAARFYKHPsirnsi 303

Qy     263 NIVVVKVLLLRDRDSGPKVTGNAALTLRNFCAWQKKLNKVSdkhPEYWDTAILFTRQDLC 322
      :| | | | | | | | | | | | | | | | | | | | | |
Db     304 SLVVVKILVIYEEQKGPEVTSNAALTLRNFCNWQKQHNSPSDRDPHYDTAILFTRQDLC 363

Qy     323 GATTCDTLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCCEVFG 382
      | : | | | | | | | | | | | | | | | | | | | | |
Db     364 GSHTCDTLGMADVGTVC DPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCASLNG 423

Qy     383 KLRANHMSPTLIQIDRANPWSACSAAIITDFLDSGHGDCLLDQPSKPISLPEDLPGASY 442
      :| :| : | : | | | | :| | | | | | | | | | | |
Db     424 VTGDShLMASMLSSLDSQPWSPCSAYMVTsFLDNHGGECLMDKPNPIKLPSDLPGTLY 483

Qy     443 TLSQQCELAfGVGSKPCP-YMQYCTKLWCTGKAKGMVCQTRHFPWADGTSCGEGKCLK 501
      :| | : | | | | | | | | | | | | | | | | | |
Db     484 DANRQCQFTFGEEKHCPDAASTCTTLWCTGTSGGLLVCQTKHFPWADGTSCGEGKWCVS 543

Qy     502 GACVERHNLNKH---RVDGSWAKWDPYGPCSRTCGGGVQLARRQCTNPTPANGGKYCEGV 558
      | | | : : | | | | | | | | | | | | | | | |
Db     544 GKCVNKTDM-KHFATPVHGSWGPWGPWGDsRTCGGGVQYTMRECDNPVPKNGGKYCEGK 602

Qy     559 RVKYRSCNLEPCPSSASGKSFREEQCEAFNGYNHSTNRLTLAVAWVPKYSVSPRKCKL 618
      | | : | | | | | | | | | | | | | | | | | |
Db     603 RVRYRSCNIEDCPDN-NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKL 661

Qy     619 ICRANGTYFYVLAPKVVDGTLCSPDSTSVCVQGKCIKAGCDGNLGSKKRFDKCGVCGGD 678
      | | | | | | | | | | | | | | | | | | | | |
Db     662 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGVCGGN 721

Qy     679 NKSCCKVTGLFTKPMHGYNFVVAIPAGASSIDIRQGYKGLIGDDNYLALKNSQGKYLLN 738
      :| | :| : | | | | | | | | | | | | | | | |
Db     722 GSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHRNQGRSRNNGSFLAIRAADGTyILN 781

Qy     739 GHFVVSAVERDLVVKGSLLRYSGTGTAVESLQASRPILPTEVLSVGKMTPPRVRYSF 798
      | : | : | | | | | | | | | | | | | | | |
Db     782 GNFTLTSTLEQDLTYKGTVLRYSGSSAALERIRSFSPLEPLTIQVLMVGHALRPKIKFTY 841

Qy     799 YLPKEPREDKSSH PKDPRGPSVLHNSVLsLSNQVEQPDDRPPARWVAGSWGPCSASCGSG 858
      : : | : | | | | | | | | | | | | | | |
Db     842 FMKKKTES-----FNAIPTFS-----EWVIEEWGECskTCGSG 874

Qy     859 LQKRAVDRCGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCSKSCGRGFQR 915
      | : | | | | | | | | | | | | | | | | | |
Db     875 WQRRVVQCRDINGH---PASECAKEVKPASTRPCADLPCPHWQVGDWSPCSKTCGKGYKK 931
```

Qy 916 RSLKCVGHGGRLRLARDQCNLHRKPQE-LDFCVLRPC 950  
|:|||| | | :|: : |: :||: :||| | |  
Db 932 RTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 967

Search completed: August 5, 2006, 00:29:46  
Job time : 190 secs

## SCORE Search Results Details for Application 10763210 and Search Result us-10-763-210-1.rai.

Score Home Page [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ](#) [Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10763210 and Search Result us-10-763-210-1.rai.  
[start](#)

[Go Back to previous](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2006, 00:14:04 ; Search time 52 Seconds  
(without alignments)  
1599.117 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H\_COMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	5164	100.0		950	2	US-10-009-332-1	Sequence 1, Appli
2	2482.5	48.1		950	2	US-09-321-987B-4	Sequence 4, Appli
3	2480.5	48.0		949	2	US-09-568-559-2	Sequence 2, Appli
4	2480.5	48.0		967	2	US-09-130-491-2	Sequence 2, Appli
5	2274	44.0		727	2	US-09-445-023A-1	Sequence 1, Appli

6	2273	44.0	727	2	US-09-445-023A-12	Sequence 12, Appl
7	2135.5	41.4	890	2	US-09-949-002-394	Sequence 394, App
8	2132.5	41.3	887	2	US-09-949-002-534	Sequence 534, App
9	2113	40.9	905	2	US-09-369-364A-9	Sequence 9, Appli
10	1917	37.1	837	2	US-09-122-126B-2	Sequence 2, Appli
11	1917	37.1	837	2	US-09-634-286A-2	Sequence 2, Appli
12	1917	37.1	837	2	US-10-247-685-2	Sequence 2, Appli
13	1917	37.1	849	2	US-09-949-002-564	Sequence 564, App
14	1916	37.1	837	2	US-10-012-231A-317	Sequence 317, App
15	1916	37.1	837	2	US-10-015-389A-317	Sequence 317, App
16	1916	37.1	837	2	US-10-006-768A-317	Sequence 317, App
17	1916	37.1	837	2	US-10-015-671A-317	Sequence 317, App
18	1916	37.1	837	2	US-10-015-393A-317	Sequence 317, App
19	1916	37.1	837	2	US-10-011-833A-317	Sequence 317, App
20	1916	37.1	837	2	US-10-006-041A-317	Sequence 317, App
21	1916	37.1	837	2	US-10-012-064A-317	Sequence 317, App
22	1916	37.1	837	2	US-10-015-392A-317	Sequence 317, App
23	1916	37.1	837	3	US-10-011-795B-317	Sequence 317, App
24	1916	37.1	837	3	US-10-015-386A-317	Sequence 317, App
25	1916	37.1	837	3	US-10-012-121A-317	Sequence 317, App
26	1916	37.1	837	3	US-10-006-485A-317	Sequence 317, App
27	1916	37.1	837	3	US-10-006-746A-317	Sequence 317, App
28	1916	37.1	837	3	US-10-012-752A-317	Sequence 317, App
29	1916	37.1	837	3	US-10-017-253A-317	Sequence 317, App
30	1916	37.1	837	3	US-10-015-519A-317	Sequence 317, App
31	1916	37.1	837	3	US-10-015-715A-317	Sequence 317, App
32	1916	37.1	837	3	US-10-007-236A-317	Sequence 317, App
33	1915	37.1	837	2	US-09-949-002-380	Sequence 380, App
34	1900	36.8	608	2	US-09-130-491-13	Sequence 13, Appl
35	1854	35.9	930	2	US-09-369-364A-2	Sequence 2, Appli
36	1848.5	35.8	1882	2	US-09-369-364A-13	Sequence 13, Appl
37	1847.5	35.8	930	2	US-09-122-126B-15	Sequence 15, Appl
38	1847.5	35.8	930	2	US-09-634-286A-15	Sequence 15, Appl
39	1847.5	35.8	930	2	US-10-247-685-15	Sequence 15, Appl
40	1798.5	34.8	874	2	US-09-369-364A-15	Sequence 15, Appl
41	1607	31.1	551	2	US-09-130-491-16	Sequence 16, Appl
42	1593.5	30.9	2150	2	US-09-321-987B-2	Sequence 2, Appli
43	1585	30.7	2165	2	US-09-800-729-155	Sequence 155, App
44	1304	25.3	1082	2	US-10-363-937-17	Sequence 17, Appl
45	1302	25.2	1104	2	US-09-981-953A-4	Sequence 4, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-009-332-1

; Sequence 1, Application US/10009332

; Patent No. 6716613

; GENERAL INFORMATION:

; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

; APPLICANT: Kazusa DNA Research Institute

; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGRECANASE ACTIVITY

; FILE REFERENCE: Q67541

; CURRENT APPLICATION NUMBER: US/10/009,332

; CURRENT FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: JPA Hei 11-321740

; PRIOR FILING DATE: 1999-11-11

; PRIOR APPLICATION NUMBER: JPA 2000-144020

; PRIOR FILING DATE: 2000-05-16

; NUMBER OF SEQ ID NOS: 35



This page gives you Search Results detail for the Application 10763210 and Search Result us-10-763-210-1.rapbn.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2006, 00:26:50 ; Search time 36 Seconds  
(without alignments)  
1765.787 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 236815 seqs, 66914042 residues

Total number of hits satisfying chosen parameters: 236815

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA\_New:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1916	37.1	837	6	US-10-196-749-352
2	1183	22.9	1201	7	US-11-300-928-40
3	551.5	10.7	525	6	US-10-196-749-436
4	349.5	6.8	813	6	US-10-196-749-466
5	317	6.1	737	7	US-11-193-353-4
6	314	6.1	735	6	US-10-196-749-88
7	314	6.1	909	7	US-11-226-554-80
8	314	6.1	909	7	US-11-248-718-80
9	276.5	5.4	5738	6	US-10-505-928-150
10	274	5.3	763	6	US-10-982-908-28

11	255	4.9	820	6	US-10-982-908-26	Sequence 26, Appl
12	254.5	4.9	5635	6	US-10-766-760-2	Sequence 2, Appli
13	237	4.6	832	6	US-10-505-928-491	Sequence 491, App
14	214	4.1	734	7	US-11-238-282-19	Sequence 19, Appl
15	214	4.1	734	7	US-11-300-928-11	Sequence 11, Appl
16	173	3.4	765	7	US-11-165-586-36	Sequence 36, Appl
17	171.5	3.3	1403	7	US-11-217-997-12	Sequence 12, Appl
18	166.5	3.2	934	7	US-11-247-437-10	Sequence 10, Appl
19	166.5	3.2	934	7	US-11-033-476-5	Sequence 5, Appli
20	164	3.2	943	7	US-11-370-424-7	Sequence 7, Appli
21	162	3.1	734	7	US-11-165-586-45	Sequence 45, Appl
22	162	3.1	1398	7	US-11-217-997-4	Sequence 4, Appli
23	162	3.1	1404	7	US-11-217-997-2	Sequence 2, Appli
24	162	3.1	1547	7	US-11-217-997-22	Sequence 22, Appl
25	162	3.1	1577	7	US-11-217-997-16	Sequence 16, Appl

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2006, 00:08:44 ; Search time 47 Seconds  
(without alignments)  
1944.807 Million cell updates/sec

Title: US-10-763-210-1  
Perfect score: 5164  
Sequence: 1 MLLLGILTLAFAGRTAGGFE.....DQCNLHRKPQELDFCVLRPC 950

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2482.5	48.1	951	2	T00017	gene ADAMTS-1 prot
2	1915	37.1	837	2	T00355	hypothetical prote
3	1604	31.1	550	2	T47158	hypothetical prote
4	1585	30.7	2165	2	T21371	hypothetical prote
5	1192	23.1	1205	2	T18517	procollagen N-endo
6	681	13.2	1558	2	C89114	protein C37C3.6a [
7	681	13.2	2167	2	T34395	hypothetical prote
8	559.5	10.8	1444	2	T18856	angiogenesis inhib
9	510.5	9.9	860	2	T16892	hypothetical prote
10	493	9.5	951	2	T00260	hypothetical prote
11	450.5	8.7	957	2	T15976	hypothetical prote
12	411.5	8.0	1059	2	T22545	hypothetical prote
13	344.5	6.7	571	2	S24789	jararhagin C precu

14	337	6.5	609	2	S55270	catrocollastatin p
15	329	6.4	903	2	S60257	meltrin alpha - mo
16	323.5	6.3	617	2	S48160	metalloproteinase
17	319	6.2	826	2	A60385	monocyte surface a
18	313	6.1	549	2	S48169	metalloproteinase
19	308	6.0	789	2	S28259	androgen-regulated
20	307	5.9	411	1	HYSNFA	fibrolase (EC 3.4.
21	304	5.9	407	2	S66260	metalloproteinase
22	300.5	5.8	616	2	A55796	ecarin precursor -
23	297	5.8	480	1	A30065	trigramin precurs
24	292	5.7	478	2	JC4880	fibrinolytic metal
25	290.5	5.6	481	2	JC4342	fibrinolytic prote
26	289	5.6	414	2	S41609	atrolysin C (EC 3.
27	286.5	5.5	610	2	JC7530	vascular apoptosis
28	284	5.5	414	1	HYRSAC	atrolysin C (EC 3.
29	281.5	5.5	610	2	JC8056	halysase - Gloydiu
30	280	5.4	481	2	S43125	trimucin precursor
31	277.5	5.4	814	2	G02390	disintegrin-like m
32	273.5	5.3	478	2	A43296	atrolysin E (EC 3.
33	265	5.1	414	2	S41608	atrolysin B (EC 3.
34	265	5.1	670	2	I65967	disintegrin-like m
35	260	5.0	655	2	JC7850	disintegrin and me
36	257	5.0	776	2	S28258	androgen-regulated
37	256.5	5.0	1074	2	JC5928	semaphorin F precu
38	255.5	4.9	1584	2	T00026	brain-specific ang
39	253.5	4.9	952	2	T18900	disintegrin and me
40	251	4.9	1042	2	T26644	hypothetical prote
41	248.5	4.8	1170	1	TSHUP1	thrombospondin 1 p
42	245.5	4.8	1170	2	A40558	thrombospondin 1 p
43	244.5	4.7	419	2	A59414	metalloproteinase
44	243	4.7	478	2	JQ1301	hemorrhagic protei
45	241.5	4.7	484	2	JC8020	metalloproteinase-

#### ALIGNMENTS

##### RESULT 1

T00017

gene ADAMTS-1 protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 15-Mar-2004

C;Accession: T00017

R;Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.

Genomics 46, 466-471, 1997

A;Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.

A;Reference number: Z14055; MUID:98110583; PMID:9441751

A;Accession: T00017

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-951 <KUN>

A;Cross-references: UNIPARC:UPI0000049F23; EMBL:AB001735; NID:g2809056;

PIDN:BAA24501.1; PID:g2809057

A;Experimental source: strain 129SVJ

C;Genetics:

A;Gene: ADAMTS-1

> gi|12644483|sp|Q9UHI8|ATS1\_HUMAN **G** ADAMTS-1 precursor (A disintegrin and metalloproteinase thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1)  
Length=967

Score = 471 bits (1213), Expect = 1e-132, Method: Composition-based stats.  
Identities = 245/387 (63%), Positives = 297/387 (76%), Gaps = 6/387 (1%)

Query 1 RAKRFVSIPRYVETLVVADESMVKFHGADLEHylltllataarlyrHPSILNPINIvvvk 60  
R KRFVS RYVET++VAD+SM +FHG+ L+HYLLTL + AARLY+HPSI N +++VVVK  
Sbjct 249 RKKRFVSSHRYVETMLVADQSMAEFHGSGLKHYLLTLFSVAARLYKHPSIRNSVSLVVVK 308

Query 61 vlllRDRDSGPKVTGNAALTLRNFCAWQKKLNKVS DKHPEYWDTAILFTRQDLCGATTCD 120  
+L++ D GP+VT NAALTLRNFC WQK+ N SD+ E++DTAILFTRQDLCG+ TCD  
Sbjct 309 ILVIHDEQKGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQTCD 368

Query 121 TLGMADVGTMCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHDNVKVCEEVFGKLRANH 180  
TLGMADVGT+CDP RSCSVIEDDGL +AFTTAHELGHVFNMPHD+ K C + G + +H  
Sbjct 369 TLGMADVGTVCDP SRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASLNGVNQDSH 428

Query 181 MMSPTLIQIDRANPWSACSAAIITDFLD SGHGDCLLDQPSKPISLPEDLP GASYTLSQQC 240  
MM+ L +D + PWS CSA +IT FLD+GHG+CL+D+P PI LP DLPG SY ++QC  
Sbjct 429 MMASMLS NLDHSQPWSPCSAYMITSFLDNHGGECLMDKPQNPIQLPGDLP GTSYDANRQC 488

Query 241 ELAFGVGSKPCP-YMQYCTKLWCTGKAKGQMV CQTRHFPWADGTSCGEGKLCCLKGACVER 299  
+ FG SK CP C+ LWCTG + G +VCQT+HFPWADGTSCGEGK C+ G CV +  
Sbjct 489 QFTFGEDSKHCPDAASTCSTLWCTGTSGGVLCVQTKHFPWADGTSCGEGKWCINGKCVNK 548

Query 300 HNLNKH---RVDGSWAKWDPYGPCSRTC GGGVQLARRQCTNPTPANGGKYCEGVRVKYRS 356  
+ KH GSW W P+G CSRTC GGGVQ R+C NP P NGGKYCEG RV+YRS  
Sbjct 549 TD-RKHFDTPFHGSGWGMWGPWGDCSRTC GGGVQYTMRECDNPVPKNGGKYCEGKRVYRS 607

Query 357 CNLEPCPSSASGKSFREEQCEAFNGYN 383  
CNLE CP + +GK+FREEQCEA N ++  
Sbjct 608 CNLEDCPDN-NGKTFREEQCEAHNEFS 633